



JOTA
Journal of Tax Administration

Volume 2, Issue 2

October 2016

CONTENTS

ARTICLES

COLLABORATIVE TAX REGULATION: CAN CONSUMERS BE ENGAGED AS ‘PARTNERS’ IN THE REGULATORY CRAFT?

Karen Boll 4

EMPIRICAL EVIDENCE ON COOPERATION BETWEEN SUB-CENTRAL TAX ADMINISTRATIONS

José María Durán-Cabré, Alejandro Esteller-Moré and Luca Salvadori 24

AN EXPLORATORY INVESTIGATION OF EXTRINSIC AND INTRINSIC MOTIVATIONS IN TAX AMNESTY DECISION-MAKING

Jonathan Farrar, Cass Hausserman 47

A GLOBAL FRAMEWORK ON THE FORMULATION AND IMPLEMENTATION OF INTERNATIONAL TAXATION STANDARDS

Satoru Araki..... 67

COMMENT ON FEIGE’S PAPER “REFLECTIONS ON THE MEANING AND MEASUREMENT OF UNOBSERVED ECONOMIES: WHAT DO WE REALLY KNOW ABOUT THE ‘SHADOW ECONOMY’?”

Friedrich Schneider..... 82

PROFESSOR SCHNEIDER’S SHADOW ECONOMY (SSE): WHAT DO WE REALLY KNOW? A REJOINDER

Edgar L. Feige..... 93

REFLECTIONS ON THE MEANING AND MEASUREMENT OF UNOBSERVED ECONOMIES: AN EDITORIAL COMMENT.

Nigar Hashimzade, Chris Heady..... 108

REVIEWS

THE INTERNATIONAL CONFERENCE ON TAX ADMINISTRATION

Simon James 109

REVIEW OF RECENT LITERATURE

Nigar Hashimzade, Lynne Oats..... 120

EDITORIAL BOARD

Managing Editors

Nigar Hashimzade, Professor of Economics, Durham University.

Chris Heady, Professor of Economics, University of Kent.

Lynne Oats, Professor of Taxation and Accounting, University of Exeter.

Editorial Advisory Panel

Judith Freedman, Professor of Taxation Law, University of Oxford.

Gareth Myles, Professor of Economics, University of Exeter.

Joel Slemrod, Paul W. McCracken Collegiate Professor of Business Economics and Public Policy at the Ross School of Business, and Professor in the Department of Economics at the University of Michigan.

Editorial Board

James Alm, Professor of Economics, Tulane University.

Peter Birch Sørensen, Professor of Economics, University of Copenhagen.

Richard M. Bird, Professor Emeritus, University of Toronto.

Rebecca Boden, Professor of Critical Management, University of Roehampton.

Valerie Braithwaite, Professor, Regulatory Institutions Network, Australian National University.

Allison Christians, H. Heward Stikeman Chair in Tax Law, McGill University.

David Duff, Professor, Peter A. Allard School of Law, University of British Columbia.

Chris Evans, Professor of Taxation, University of New South Wales.

Anne Fairpo, Barrister, Temple Tax Chambers.

Miguel Fonseca, Associate Professor of Economics, University of Exeter.

Jane Frecknall-Hughes, Professor of Accounting and Taxation, University of Hull.

Norman Gemmell, Professor of Public Finance, Victoria University of Wellington.

Hans Gribnau, Professor of Law, University of Tilburg.

Ann Hansford, Senior Lecturer in Taxation, University of Exeter.

John Hasseldine, Professor of Accounting, University of New Hampshire.

Kristin Hickman, Harlan Albert Rogers Professor of Law, University of Minnesota.

Kevin Holland, Professor of Accounting and Taxation, University of Cardiff.

Simon James, Associate Professor of Economics, University of Exeter.

Erich Kirchler, Professor of Psychology, University of Vienna.

Christos Kotsogiannis, Professor of Economics, University of Exeter.

Emer Mulligan, Lecturer in Taxation, National University of Ireland, Galway.

David Salter, Senior Associate Fellow, University of Warwick.

Chantal Stebbings, Professor of Law and Legal History, University of Exeter.

Penelope Tuck, Professor of Accounting, Public Finance and Policy, University of Birmingham.

John Vella, Senior Research Fellow, Oxford Centre for Business Taxation.

ABOUT THE JOURNAL

The Journal of Tax Administration is a peer-reviewed, open access journal concerned with all aspects of tax administration. Initiated in 2014, it is a joint venture between the University of Exeter and the Chartered Institute of Taxation.

JOTA provides an interdisciplinary forum for research on all aspects of tax administration. Research in this area is currently widely dispersed across a range of outlets, making it difficult to keep abreast of. Tax administration can also be approached from a variety of perspectives including, but not limited to, accounting, economics, psychology, sociology and law. JOTA seeks to bring together these disparate perspectives within a single source to engender more nuanced debate about this significant aspect of socio-economic relations. Submissions are welcome from both researchers and practitioners on tax compliance, tax authority organisation and functioning, comparative tax administration and global developments.

The editorial team welcomes a wide variety of methodological approaches, including analytical modelling, archival, experimental, survey, qualitative and descriptive approaches. Submitted papers are subjected to a rigorous blind peer review process.

SUBMISSION OF PAPERS

In preparing papers for submission to the journal, authors are requested to bear in mind the diverse readership, which includes academics from a wide range of disciplinary backgrounds, tax policy makers and administrators, and tax practitioners. Technical and methodological discussion should be tailored accordingly and lengthy mathematical derivations, if any, should be located in appendices.

MESSAGE FROM THE CHARTERED INSTITUTE OF TAXATION

The Chartered Institute of Taxation is an education charity with a remit to advance public education in, and the promotion of, the study of the administration and practice of taxation. Although we are best known for the professional examinations for our members, we have also supported the academic study of taxation for many years and are pleased to widen that support with our involvement with this journal.

WEBSITE

The Journal of Tax Administration website can be found here: www.jota.website

SOCIAL MEDIA

We also have a Twitter account: <https://twitter.com/jotajournal>

EDITORIAL NOTE

We are pleased to present the second 2016 issue of the Journal of Tax Administration which once again contains a geographically dispersed and methodologically diverse set of papers. We are grateful to all contributors to this issue, both authors and reviewers.

In the first paper, Karen Boll explores a new form of regulation taking place in Denmark that encourages consumers to only purchase services from tax compliant businesses. This specific case of innovation in tax administration is linked to broader theoretical literature on collaborative and interactive governance, and draws on ethnographic fieldwork conducted with Danish Customs and Tax Administration employees working on a project relating to the cleaning sector.

The second paper, by José María Durán-Cabré and colleagues deals with an under-researched area of public finance: that of cooperation between tax authorities at sub-central level. The setting is Spanish regional tax authorities and the distribution of wealth tax revenues among them which necessitates cooperative behaviour. The authors find that once regional tax administrations become aware of the potential benefits of cooperation, engagement in it is subsequently maintained.

The third paper, by Jonathan Farrar and Cass Hausserman from Canada, reports an exploratory quasi-experimental investigation of tax amnesties that considers extrinsic and intrinsic motivations for decisions, the latter being an under-researched area generally, but especially in the specific context of tax amnesties. The authors use conjoint analysis with a supplemental analysis, and find that the desire to avoid a penalty was the most important extrinsic motive, while responsibility for paying taxes owed was the most important intrinsic motivation.

The fourth paper is by Satoru Araki from Japan who, based on substantial practical experience as a tax administrator, proposes a nascent global architecture of international tax standards that importantly embraces developing countries. This architecture is evolving and the author suggests that it requires a solid network of regionally based frameworks to enable a truly global reach.

A new feature in this issue of JOTA is an exchange between two scholars who research estimations of the extent of the shadow economy. Issue 1 of Volume 2 of JOTA, published earlier this year, contained a paper by Professor Feige, in which he expressed concern about the work of a fellow scholar, Professor Schneider. We took the decision to allow each of them to write a follow-up in this issue of the journal in order to foster methodological debate. Professor Schneider provides comments on Professor Feige's previously published work and Professor Feige provides a rejoinder in this issue.

Finally, this issue contains two review papers. The first, by Simon James, charts the history of the Australian Tax Administration Conference and outlines the papers presented at the 2016 conference that was held in Sydney from 31 March to 1 April 2016. The second is a review of recent literature that canvasses tax administration-related publications from a variety of sources during the first half of 2016. We hope readers find this useful and we welcome suggestions for inclusion in future reviews.

Lynne Oats (on behalf of the Managing Editors)

COLLABORATIVE TAX REGULATION: CAN CONSUMERS BE ENGAGED AS 'PARTNERS' IN THE REGULATORY CRAFT?

*Karen Boll*¹

Abstract

This article shows a new form of regulation within a tax administration where tax administrators abate tax evasion by nudging and motivating consumers to only purchase services from tax compliant businesses. This indirectly closes or forces tax evading businesses to change their practices, because their customer bases decline to commercially non-viable levels. The analysis is framed by public governance literature and argues that the regulation is an example of *collaborative or interactive governance*, because the tax administrators do not regulate non-compliance directly, but activate external stakeholders, i.e. the consumers, in the regulatory craft. The study is based on a qualitative methodology and draws on a unique case of regulation in the cleaning sector. This sector is at high risk of tax evasion and human exploitation of vulnerable workers operating in the informal economy. The article has implications for how tax practitioners think about collaborative and interactive regulatory initiatives. While the tax administration in the study sees the approach as effective, the analysis shows that there are a number of caveats in relation to regularity, public listing, costs and revenue focus. The article thus links a concrete case of new regulation in tax administration to broader theoretical discussions of collaboration and interactive governance within public administration, and the article problematizes this regulation. This provides a vantage point from which constructive dialogue about new regulatory practice can emerge.

INTRODUCTION

In addition to having a viable cash flow, a key premise for running a business is compliance with rules and regulations concerning business activities. Businesses that make money but do not comply with the tax law (and/or other laws) will, at some point, experience intervention by public tax administrators or other public office-holders. Commonly, tax administrators use the legal code to enforce compliance via inspections, sanctions or fines directly targeted at businesses that breach the law. In contrast to this approach, a new trend of 'cooperative' enforcement is discernible within tax administration. Working with consumers, tax administrators actively influence the cash flow of non-compliant businesses by informing customers of the non-compliant tax practices of the selling businesses. Such businesses may either close or comply, not because of direct enforcement from tax administrators, but because of declining sales to consumers who are increasingly suspicious of the businesses' integrity. Tax compliance here becomes the result of 'cooperation' between tax administrators and consumers.

Such *collaborative regulation* within public administration has been the focus of much recent research. Scholarship has theorised what occurs when government officials engage in indirect

¹ Associate Professor, Copenhagen Business School.

forms of governance and when government officials collaborate with external stakeholders to achieve their goals. Flinders has discussed distributed governance (2005); Grabosky has discussed the use of non-governmental resources to foster regulatory compliance (1995); Rhodes has discussed network governance (1997); Torfing et al. have discussed interactive governance (2012); and Alford and others have advocated the potential of co-production (Alford, 2009, 2013; Needham, 2008; Ryan, 2012; van Eijk & Steen, 2013). Within this body of scholarship, this article refers to the work of Torfing et al. and their concept of *interactive governance* (Torfing et al., 2012). This concept describes governance that occurs through a plurality of state and non-state actors (ibid. 2).

The article focusses on actual tax practices by showing how tax administrators from the Danish Customs and Tax Administration regulate tax evasion by motivating consumers to purchase services from tax compliant businesses. This indirectly closes or forces tax evading businesses to change their practices because the customer base declines to commercially non-viable levels. I argue that this regulation is an example of *collaborative* or *interactive governance* because the tax administrators activate external stakeholders (that is, the consumers) in the regulatory craft. The study employs a qualitative methodology and focusses on the regulation of contracting in the cleaning sector. The significance of the analysis is that it problematizes this new form of collaborative regulation. The tax administrators themselves see the approach as effective, yet this analysis shows that there are a number of caveats in relation to regularity, public listing, costs and revenue focus. The contribution of the article is that it links a case of enforcement in tax administration to theoretical discussions of *interactive governance* within public administration and problematizes this regulation. In this way, rigorous use of theory from public administration is connected specifically to regulation within tax administration.

The following section introduces the theoretical concepts used to study collaborative regulation. The subsequent section is a methods section that introduces the qualitative study. The body of the article presents and analyses the case. The penultimate section discusses the new regulation strategy and highlights its challenges. The final section is the conclusion, together with an outline of the article's contribution to the literature on collaborative and interactive governance within public administration.

COLLABORATIVE REGULATION

It is a common experience across different public sector institutions that traditional state-centric forms of governance are being challenged. Today, it is often not *comme-il-faut* for public service providers to rely on top-down imposition of authority - to be formalistic, rigid, closed and hierarchical. Rather, when public services are developed, these should be co-produced with users who are involved in the process (Jakobsen & Andersen, 2013). Many large-scale public sector reforms are based on thorough engagement and involvement of the affected parties who can influence the policy process bottom-up (Dunston, Lee, Boud, Brodie, & Chiarella, 2009). *User-involvement, participation, collaboration, partnerships* and *co-production* are buzzwords for today's governance. In continuation of this, Torfing et al. (2012) note that it is central to recognise that the state cannot be seen as the sole actor involved, nor can the state necessarily be seen as the most fundamental actor in governance. No single actor, they argue, can account for contemporary governance alone. Instead, contemporary governance initiatives involve government, markets

and/or some forms of *interactive governance* based on a combination of actors (ibid. 4-5). In other words, there are limits to unilateral state action (ibid. 4).

This thought is perhaps not so provocative within areas such as public health care, schooling or planning, where user-involvement and collaboration have been standard development methods for years, and where experiences have been thoroughly documented and analysed (Bovaird, 2007; Dunston et al., 2009; Fledderus, Brandsen, & Honingh, 2013; Harrits & Møller, 2013; Needham, 2008; Podger, 2012; van Eijk & Steen, 2013). Among the more coercive functions of the state, such as the police, the military and - of interest here - tax administration, *partnership*, *participation*, *co-production* and comparable strategies have had less traction.

These coercive functions of the state have regulatory and law enforcement responsibilities that they predominantly exercise themselves, and often these authorities engage with citizens and businesses who are not at all pleased about their actions. Who really likes to be arrested? Fined? Have a license revoked? Or have one's property seized? (Sparrow, 2000). It can simply be unpleasant to be forced into compliance by any of these authorities, and the citizenry and businesses may receive treatments that they have not requested, have not paid for, might not enjoy, and which they will not want to repeat. As Sparrow writes, the core of these authorities involves the imposition of duties. These authorities deliver obligations, rather than services (ibid. 2).

Although Sparrow is right about these authorities' coercive enforcement responsibilities, there has been a trend in recent years for tax administrations, in particular, to be increasingly attentive to the possibility that, in some areas of regulation, partnering with external stakeholders is a viable route to securing compliance. This approach has been blueprinted in the Organisation for Economic Cooperation and Development (OECD) report *Together for better outcomes: Engaging and Involving SME Taxpayers and Stakeholders* (2013). This report emphasises the potential for creating better tax compliance by initiating cooperation with external stakeholders and states that: "On their own [the revenue bodies] are not capable of addressing the scale of the challenge revenue bodies face, particularly in the wake of the global financial crisis" (ibid. 2).

Instead, revenue bodies increasingly need to look outside their own organisations to use the knowledge and resources of both taxpayers and other stakeholders to achieve greater tax compliance. If they do this, there is potential for improving outcomes, enhancing services and reducing costs (ibid. 3). Thus, in this way, the OECD report concurs with Torfing et al.'s description that there are 'limits to unilateral state action' (Torfing et al., 2012, p. 4) and even encourages tax administrations to engage in participation, collaboration and partnerships to achieve their goals.

INTERACTIVE GOVERNANCE

Although the OECD promotes the engagement and involvement of external stakeholders in fostering tax compliance, the OECD does not provide theoretical concepts to analyse what happens when engagement and involvement efforts are initiated in practice. On the other hand, this is what Torfing et al. do with their concept of *interactive governance*. As briefly sketched in the introduction, interactive governance describes how a plurality of actors - state actors together with non-state actors - interact to promote and achieve common objectives (Torfing et al., 2012, p. 2).

The concept stresses the element of so-called meta-governance. This means that governments play a central role in meta-governing interactive governance by means of their institutional conditions and by designing, managing and directing the interactive governance arenas (ibid. 4). Another facet of the concept is that it highlights the new roles and responsibilities of the bureaucrats and citizens/businesses involved. The bureaucrats become less like ‘hands-on’ regulators and more like managers of interactions (that is, designers of the interactive areas). In addition, the regulatees - that is, those that are being regulated - take on new roles. Previously, they were subjects who were regulated. Now they are recast as co-producers of governance and as ‘partners’. In sum, the concept emphasises a shift toward a new form of governance in which states no longer steer primarily through formal state actors, but instead move towards indirect, collaborative and interactive steering where several stakeholders are engaged. All of this changes the roles and responsibilities of the parties involved and may blur previously established distinctions between regulator and regulatee (ibid. 151).

Although this paper draws on Torfing et al.’s conceptualization of interactive governance, a disclaimer should also be provided concerning Torfing et al.’s understanding of tax administration/revenue collection. Their work does not go into detail on this area of governance but, in several instances, they use revenue collection as an example - together with foreign affairs and defence - where they believe formal and legalistic state-based steering continues to be important, because these areas are seen as core defining functions of the state (ibid. 2). They write:

Likewise, this form of governance [i.e., interactive governance] may not be appropriate for all policy areas, especially those that involve (...) the ‘defining functions’ of the state, for example, law, defence, and taxation (ibid. 4).

Although I concur that taxation is a defining function of the state, I object to the tendency to categorise tax administration as based on the traditional top-down imposition of authority and describing it as an area not appropriate for interactive forms of governance. More specifically, I view this description as an expression of a lack of knowledge of contemporary trends in tax administration. More knowledge of actual practices and procedures within tax administration is needed. As evident in this article, interactive forms of governance already exist in tax administration and the regulation developed in this area of government is more subtle than the ‘stereotypical’ perceptions expressed by Torfing et al. and partly by Sparrow.

ORGANISATIONAL ETHNOGRAPHY

In her book, *Taxation: A Fieldwork Research Handbook* (2012), Lynne Oats argues that tax is a social and institutional practice and that, often, not enough attention is paid to the important aspect of tax as an institution (ibid. 5). This social and institutional aspect of taxation is adhered to in this study. It is the tax administration - this *institution* - which is in focus. What will be presented is a detailed empirical account of the way that tax inspectors put certain tax rules and regulations into play in their governance efforts.

The method employed in the study is qualitative, and is inspired by *organisational and administrative ethnography* (Boll & Rhodes, 2015; Moeran, 2005; Neyland, 2007; Yanow, Ybema, & van Hulst, 2012; Ybema, Yanow, & Kamsteeg, 2009). This approach uses ethnography, that is, methods such as observation, participation and explorative interviewing (e.g. Hammersley

& Atkinson, 2007), to study an organisational setting. What one gains from observing work, participating in meetings and interviewing organisational members is knowledge about how 'things work around here' (Rhodes, 2011, 2015). This methodology enables knowledge about the everyday practices and experiences of the people working in these organisations. When using this method to study public administration, one gains insight into how these organisations are 'commonplaces' where the state enacts itself.

Ybema et al. suggest that the quotidian experiences of people working in these organisations may, to some, hardly seem exciting (2009, p.1). On the one hand, I can appreciate this observation. Sometimes when I enter a tax office, I question whether there is something exciting to find when I see the legendary 'paper-pushing-bureaucrat' bent over the keyboard. Yet, on the other hand, when I start to engage with the inspectors, when I hear about their work, their beliefs and practices, when I follow them on their inspections, then I always become intrigued by the complexity and multifaceted character of their work. These people often struggle to meet organisational demands, to follow new strategies, to provide sound instruction and services to citizens and to maintain their motivation for, and adherence to, the ethical standards of their work. These intriguing aspects of administrative life - in taxation - are what I aim to convey in the following analysis.

Having *organisational ethnography* as a methodological marker to steer with is challenging. As Neyland notes, ethnography is by no means a straightforward methodology. It requires a great deal of access to the field being studied, a participative role for the researchers, a great deal of time spent in the field and a great deal of researcher involvement in gathering, organising and analysing observations (2007, p. 2ff). Relating to these challenges, the first hurdle to overcome when doing organisational ethnography in tax administration is to gain access. By their nature, tax inspectors handle confidential information about tax payers. Speaking from experience, it is almost impossible to sit at the desk of a tax inspector without one's eyes gazing at memorandums, letters, notes, audit results and reports. Many of these documents have attached names, addresses and central person registration numbers. Allowing a researcher to access this 'room' requires extensive agreements about confidentiality, secrecy and anonymity. A second hurdle to overcome when seeking to conduct fieldwork is being present when tax inspectors interact with taxpayers. This may be during service encounters, random audits or actual inspections. Being present as a researcher on these occasions often requires additional consent from the citizens and additional agreements on secrecy.

Another challenge when aspiring to do organisational ethnography in a public administration setting is *time* - the informants' time. I have been studying tax administration in Denmark for a number of years (Boll, 2014, 2015). Without exception, I have met tax administrators who are conscious about how they spend their time because they are pressed for time and resources - often having stacks of cases on their desk. For these people, it is an inconvenience to have an ethnographer hanging around. Work slows down: procedures and practices have to be explained, precautions need to be taken regarding confidentiality, and interactions need to be cleared by the senior managers. It is an additional pressure to have two eyes following you. When I follow these inspectors, I sense that they try to do their work in the best possible manner - no mistakes can be made when the researcher is there.

Any academic doing empirical work is likely to face many of these challenges. However, the challenges are pronounced for organisational ethnographers because the key to this method is to stay for a long time and to ‘get close’ to the informants’ work. Due to the constraints in relation to time, access and secrecy, I have found it difficult to justify explorative fieldwork over longer periods of time. It strains the tax administrators. In contrast, I have had positive experiences conducting shorter and more focussed fieldwork. The case below exemplifies such a tightly focussed fieldwork project. Importantly, this fieldwork is not *organisational ethnography* as such - that would have required more observation and/or participation over a longer period of time. Instead, organisational ethnography is the methodological marker that the data collection is steered by. With the sources at hand, the study seeks to describe and analyse how ‘things work around here’. It seeks to convey the tax administrators’ everyday work practices.

THE DATA

In the spring of 2012, I conducted fieldwork in a tax unit in the Danish Customs and Tax Administration. This unit was responsible for the regulation of contracting in the cleaning sector. The unit had its operating core in a tax office north of Copenhagen, Denmark. The staffing of such a unit - or ‘project’, as it is colloquially referred to - typically includes a *project owner*, who has overall responsibility for the project. This person often oversees several projects/initiatives. A *project leader* coordinates and manages the project on a daily basis. This person typically runs only one project at a time and is formally responsible for internal documentation, such as the project description, evaluations and staffing. As some projects/initiatives are run simultaneously across the entire country, some of the project leaders are also called *regional coordinators*, indicating that they coordinate and manage a larger regional effort. Finally, there are the *case workers*. These individuals are responsible for the actual work. Some of the case workers are allocated to a project/initiative full-time, whereas others work on two or three projects.

During the fieldwork, I was permitted to interview a broad spectrum of employees. This resulted in a total of 18 interviews. Two interviewees were conducted with *project owners*, three with *project leaders/regional coordinators* and 13 with *case workers*. All interviews were explorative in their nature. Some of the interviews focussed directly on the regulation of the cleaning sector, while some of the interviews had a basis in other related areas of regulation. The aim was to get the involved staff to discuss their work and its history in the organisation, and to explain what the challenging, rewarding and joyful aspects of what they were doing are. The interviews were conducted at the tax office, and all interviews were recorded and transcribed.

In addition to the interviews, I was also given a number of internal administration documents. These documents are listed in Figure 1. The documents contain confidential information about the Tax Administration’s work procedures. To access them, I signed a confidentiality agreement, which was negotiated so that I could describe cases from this material, as long as the information was anonymised. Furthermore, I obtained a number of publicly available documents relating to the work in focus. This material includes four questions from the Danish Parliament to the Tax Administration, 14 newspaper articles about the work in focus, and a press release from the Tax Administration.

Figure 1: Overview of internal documents. (N.B. I cite these documents continuously in the analysis).

Document	Project: Regulating the cleaning sector
Internal document	Protocol for work
Internal document	Project description
Internal document	PowerPoint – final evaluation
Internal document	Status – to the Minister

Together, this material has allowed me to describe and analyse how the tax administrators regulate the contracting of cleaning services. Although the case is based in a specific tax agency and national context, I believe that the mechanism of collaboration in regulation illustrates a trend that is recognisable for other tax administrations responsible for regulation. Hopefully, the case can provide food for thought, despite its national specificity. The data also included classic observational fieldwork. This has been reported separately in an article in *Journal of Organizational Ethnography* (Boll, 2015).

REGULATING THE CLEANING SECTOR

In Denmark, as in many other Western countries, cleaning is often conducted by migrant workers. For many of these workers, a cleaning job is attractive because it requires few language skills, is unskilled and provides a stable source of income which can (if the cleaners are not exploited) be considerable compared to salaries in their home countries. Although attractive as a job, the cleaning sector is also a ‘Wild West’. In Denmark, the Tax Administration’s random inspection of cleaners’ working conditions shows that three out of five cleaners are paid ‘cash in hand’, work too many hours, or do not possess a work permit, contract, or residence permit (project description). There are severe problems with workers regularly being exploited by middlemen. These middlemen are organised as subcontractors and hire the cleaners under illicit conditions. The taxation obligations of these subcontractors are often sidestepped (project description). The cleaning sector is, therefore, an area ripe for both human exploitation and tax evasion.

In the interviews, the case workers explain that the cleaning they have in focus for their regulation is cleaning done at public hospitals, schools, universities or large private corporations. In most cases, these purchasers of cleaning services have signed a contract with a cleaning company and have contracted in good faith. However, the cleaning companies that sign the contracts often do not employ cleaners themselves. Instead, they hire sub-contractors, who again hire sub-contractors, who then provide the cleaners from their networks. It is not uncommon for cleaners from these networks to be migrant workers who do not have work or residence permits. These workers operate in the informal part of society and the economy. In the chain of middlemen who organise the cleaning, consideration for wages, taxes, contracts and work permits disappear - also because such formal documentation simply cannot be presented. As a result, the cleaners who show up to do the cleaning at the schools, hospitals or supermarkets have been given official clothing, and can often provide papers, contracts or IDs, but most of this documentation is fake. The case workers explain that the purchasers of these cleaning services have no immediate reasons for suspicion; a contract

has been signed with a formal cleaning company and the cleaning is being done. The cleaners do not disclose that they have been hired by more or less 'shady' sub-contractors, who pay them 'cash-in-hand' and who often do not adhere to their formal taxation obligations. The cleaners keep quiet about their employment conditions, fearing to lose their jobs.

The classic strategy for regulating tax evasion in this set-up has been to target the sub-contractors by approaching the cleaners. The cleaners are often exploited and are in vulnerable or weak positions. Their positions stand in contrast to that of the subcontractors, who often deliberately take advantage of the cleaners' weaker situations. The subcontractors have a lucrative niche, because they function as the middlemen between the *cleaners*, who are eager to work but are outside the formal labour market, and the *cleaning companies* and their *consumers*. The cleaning companies want the cleaning to be done as cheaply as possible to attract customers and, therefore, they hire sub-contractors who can provide cheap cleaners. To stop this continuous supply of cleaners from 'shady' subcontractors, the case workers from the Tax Administration will show up at the locations where the cleaning is being done and will try to find the responsible subcontractors. As a result, the case workers will find cleaners with inadequate or fake papers, identify the subcontractors who have hired them, and reconstruct the earnings of these subcontractors. Finally, the subcontractors are charged for tax evasion because they pay their cleaners 'cash in hand' and do not meet their taxation obligations.

The case workers explain that they have extensive experience and are skilled in conducting such reactive investigative work. However, they also have a clear sense of the pointlessness of this work because the subcontractors do not react to the normal sanctions of penalties or forced closure of their businesses. Instead, the case workers find that when the subcontractors have been charged, they flee the country or they simply close their business to start a new one in another name. In addition, many of the subcontractors engage in money laundering, whereby large sums acquired by tax evasion in Denmark are moved into the legal economies of other countries by means of international trade. Hence, when Unger, in her research on money laundering (2009, 2013), concludes that practice is increasing rather than declining, it is tax evasion such as this in the cleaning industry that adds to the overall increase in money laundering. Looking at the Danish context, the result is that the subcontractors have no savings or assets with which to pay their tax debts.

ENGAGING CONSUMERS IN SECURING COMPLIANCE

Faced with this situation, the Danish Customs and Tax Administration decided to rethink its regulation. The project owner and leader in charge explained that they wished to change the regulation by influencing the *consumers* of the cleaning services. If those that purchased the cleaning services could be engaged in securing proper contracting and tax compliance, then much would be gained in preventing subsequent tax evasion. The tax administrators wanted to make the consumers aware of the tax evasion and human exploitation that they potentially supported when they contracted overly cheap cleaning services. The basic idea was that *if* these consumers preemptively stopped purchasing services from cleaning companies using 'shady' sub-contractors, *then* the means of existence of these businesses would diminish, thus forcing them to either comply or close. For another, yet related, use of 'public disclosure' within tax administration, see Boll and Tell (2016).

The central question for the tax administrators was how to motivate the consumers to take such pre-emptive action. The case workers explained that, for many purchasers of cleaning services, this is an expenditure that they are more than willing to cut to a minimum. One of the case workers illustrates this point by referring to a conversation she had had with a head of a public school. The head of the school explained that the school had recently saved money on their cleaning contract: the cost of their old contract was approximately 7 million DKK [1 million Euros], whereas the cost of the new contract was 5.5 million DKK [700,000 Euros]. The head of the school did say that she thought that this price was suspiciously low, but the price was provided as part of a public competitive bidding process and she needed to take the cheapest offer. Furthermore, she explained that she intended to keep an eye on the cleaning to ensure that it was done properly. In most cases, this meant inspecting whether floors were being washed, whether toilets were clean, or whether too much dust had accumulated in the corners. The *cleaning* was the focus. As such, the head of school, or any other person responsible for the cleaning, would typically look at the services provided, not at whether the cleaners receive ‘cash in hand’, have work permits, or whether the person is registered in the Tax Administration’s database as a wage earner. These administrative issues are assumed to be compliant. Yet, it is precisely in relation to these, and in relation to (under)payment, that problems arise. The consumers are less likely to pay attention to this because it involves ‘mundane’ bureaucratic issues, rather than direct delivery of services.

To inform current and potential purchasers of cleaning services about these problems and to motivate them to be more cautious when contracting them, the Tax Administration began different initiatives. Firstly, guidance meetings were set up where the case workers visited selected larger consumers to inform them about the exploitation of cleaners and how they could prevent tax evasion. This information suggested, for instance, that the consumers refuse to accept the use of subcontractors, that they routinely check the cleaners’ IDs, and that they demand that wages be paid into bank accounts. If the consumers insist on these precautions, the risk of tax evasion will be reduced. The guidance meeting strategy that targeted individual larger consumers was run together with a proactive media strategy aimed at raising public awareness about the challenges in the cleaning sectors (see also Boll, 2016). This latter strategy reflects some of the more problematic elements of the ‘collaborative’ regulation.

The two initiatives described above are types of ‘street-level’ (Lipsky, 2010 [1980]) work performed by local case workers. During my fieldwork, I was also invited to observe one meeting focussing on how senior tax administrators tried to influence the revision of the public procurement rules. During this meeting (and many other meetings), the senior tax administrators sought to include criteria based on tax compliance into the public procurement regime. I mention this to highlight the fact that the challenges of contracting in the cleaning sector were approached from various vantage points and not only from the ‘street-level’, which is the focus in this article. The reason these procurement negotiations are not included in this analysis is that I was denied access to the negotiations.

THE PROACTIVE MEDIA STRATEGY

Looking at the central elements of the ‘street-level’ initiatives, the case workers explain that their proactive media strategy relied on two elements. Initially, the case workers were to conduct a number of unannounced inspections of the cleaning at media-sensitive purchasers’ properties. These purchasers were chosen because they showed corporate social responsibility (CSR) and

cared for their public reputation. Next, if non-compliant cleaning was found, the tax administration would produce press releases informing the public about this, with an emphasis on the tax evasion. The concerned consumers, who are *not* formally responsible for the tax evasion (or for the potential human exploitation), as they have contracted in good faith, could choose to be anonymous or be listed by name in the press releases. Most purchasers chose anonymity. This was because, even though the evasion and exploitation was connected to the cleaning companies and the subcontractors who are legally responsible, the place of the evasion and exploitation was the corporate site. Simply listing this place in a press release connects this site to the problems. This does not make for flattering press coverage for the CSR-sensitive purchasers.

This press release strategy produced a number of stories about how various consumers, such as public institutions and private corporations (not listed by name) had, without knowing it, had cleaners on their property who had neither residence nor work permits, and who were underpaid and had been employed without proper contracts. The intention from the Tax Administration was that these stories would create public awareness of the problem, and encourage other purchasers of cleaning services to pre-emptively check their own contracts and their own cleaners' conditions of work. The idea was simply to distribute information.

To understand the scope of (and *spin-off* from) this strategy, one press release should be discussed in more detail. Following an unannounced inspection in January 2011, the Tax Administration sent out a press release titled: "Cleaners in newspaper group arrested" (SKAT, 2011). The press release stated that an inspection had revealed a new example of problems with subcontractors in the cleaning sector. It described how, together with the police, the Tax Administration had inspected the cleaning services in the *media industry* and that here, in one of the inspected *newspaper groups*, four cleaners had been arrested by the police. Two of these cleaners had subsequently been deported; a third cleaner had already been refused entry to the country and had been deported again; and the last cleaner had been charged for presenting false ID papers. The press release stated that the inspection showed that the cleaning company responsible for the cleaning at the newspaper group had no control of its subcontractor and that the subcontractor would be assessed with financial penalties for having employed illegal workers and for not meeting his taxation obligations.

The press release did not list the name of the purchaser, the cleaning company or the subcontractor. All were anonymised, like most of the other press releases sent out. Nevertheless, it only took a few hours before the internet-based newspaper *Journalisten.dk* publicly listed the names. It was announced that the site of the cleaning was Berlingske. Berlingske is one of the oldest and largest media groups in Denmark and is responsible for the production of several of Denmark's leading print and online newspapers. It was also reported that the cleaning company in charge was Forenade Service. Hence, within 24 hours, the purchaser of the cleaning services, Berlingske, was publicly identified in the media and linked to a story of police investigation, arrests, illegal work and tax evasion. This connection appears despite Berlingske having contracted in good faith and not having been responsible for the compliance of its cleaning company's subcontractors. Because of this stir, Berlingske was urged to take action. Less than two weeks later, Berlingske announced that it would change its cleaning service supplier. All the Berlingske offices and locations that Forenade Service was previously responsible for servicing would have a new provider.

When interviewed about this incident, the case workers described it as a success. The Tax Administration's aim with the proactive media strategy was to publish stories of how respectable and CSR-sensitive consumers unknowingly have illegal cleaners working at their properties; cleaners who were in a position to be deported from the country, who were being paid 'cash in hand', and who were, most likely, also being underpaid. As a direct result of the fact that Berlingske had its name listed publicly (revealed by independent journalists' research) and was connected to non-compliant cleaning at its sites, this purchaser decided to terminate its collaboration with its cleaning company and to hire a new, compliant provider. The case workers explained that press releases such as the one focussing on Berlingske help to raise awareness of the problems in the cleaning sector and, importantly, encourage other consumers to pre-emptively monitor their own contracts and the conditions of their own cleaners. In short, purchasers of cleaning services should see the story, identify with the problem, and take action in their own organisation to prevent a similar press story emerging based on cleaning at their properties.

INTERACTIVE GOVERNANCE - AND ITS CHALLENGES

Recalling Torfing et al.'s concept of *interactive governance*, I will argue that this concept can characterise the regulatory work taking place. In the regulatory work described above, the purchasers of the cleaning services (who are actors outside the Tax Administration) are prompted to take different actions to ensure tax compliance. These consumers may refuse to accept that their cleaning companies use subcontractors, they may start to routinely check the cleaners' IDs, or they may demand that wages be paid into bank accounts. All of these actions are prompted by the case workers either during the guidance meetings or through the proactive media strategy. If the purchasers/consumers take any of these actions, it most likely results in a situation where tax evading cleaning companies and subcontractors find it more difficult to operate, because they rely on false IDs and 'cash in hand' payments. Thus, a consequence is that the non-compliant actors either close or shift to more compliant tax practices in order to survive. The case workers describe that they find this regulatory approach compelling, because non-compliant cleaning companies and subcontractors are out-played - not by the tax administration's own direct enforcement, but by the consumers' actions. The point is that in this regulatory set-up, the consumers/purchasers help the tax administration to accomplish its aims. This set-up is a collaborative or interactive form of regulation, because state and non-state actors collaborate to achieve tax compliance.

In this regulatory set-up, the 'collaborators' take on new roles and responsibilities. With this initiative, we see that the Tax Administration moves from a situation where it used to regulate primarily through its own reactive audits, to a situation where the regulation is indirect and achieved with the activation of external stakeholders/consumers. The Tax Administration takes on a new role because it prompts, motivates or nudges these consumers to 'play out' the dodgy suppliers. In this way, the Tax Administration engages in what Torfing et al. call meta-governance, because the case workers become less like 'hands-on' regulators and more like managers of interactions, that is, designers of the interactive areas between purchasers and providers of services. As Torfing et al. describe it, "Public administrators are (...) recast as managers of interaction" (ibid. 156). What is fascinating in this regulatory set-up is that a tax-evading subcontractor may not even notice the role and work of the case workers. The subcontractor may only notice that it cannot contract with its usual cleaning company unless it pays salaries to bank accounts. It may choose to do so or choose to offer its services to another cleaning company with less strict requirements.

The purchasers of cleaning services are also invited to engage in new roles. Previously they were not aware of, or attentive to the fact, that it is worthwhile to check up on cleaning contracts and the mundane bureaucratic aspects of their cleaning. Previously, they would only inspect the 'quality' of the cleaning services delivered, but now they also need to inspect the 'working conditions'. By doing this, they become an extension of the Tax Administration because they focus directly on regulatory and tax compliance problems. Instead of being 'passive' purchasers of services, they become 'active' co-producers of governance.

As indicated in the analysis, the collaborative regulation is seen as attractive from the Tax Administration's point of view; it is considered a *smart* and *innovative* form of regulation. Although this is positive from the Tax Administration's point of view, the theory on interactive and collaborative regulation highlights that a number of challenges may appear in these kinds of regulatory arrangements. The remaining part of this section will focus on four challenges related to 1) the regularity of the regulation; 2) public listing as a tool for 'pressure'; 3) the movement of compliance costs from the Tax Administration to its 'collaborator'; and 4) the potential lost revenue focus, as the Tax Administration becomes more of a *social actor* than a *tax collector*. All of these themes problematize the engagement of consumers in the regulatory craft.

CHALLENGES WITH REGULARITY

When engaging in this form of collaborative regulation, a core task of the case workers is to influence the interactions between, on the one hand, the consumers and, on the other hand, the cleaning companies and the subcontractors. Information about tax evasion and the exploitation of cleaners becomes a type of a 'hard currency' that the tax inspectors control the flow of, and they 'spin' it in various ways to ensure that the information is interpreted by the receivers to change the market of these services in favourable ways (e.g. Torfing et al., 2012, p. 220). Hence, as described above, the case workers design a set-up in which the purchasers are engaged in enforcement and their 'power to act' is steered by the case workers, because these supply information that enables the purchasers/contractors of the cleaning services to act.

An obvious challenge in this is that the Tax Administration can try to steer what the purchasers do with the information they get, but they cannot fully control this. For instance, some consumers of cleaning services might not want to, or cannot, act as policemen in checking the validity of IDs, or might not want to include clauses in their contracts such as wage payments being paid out to bank accounts. These requirements might make the contracts with the cleaning companies more expensive. The purchasers might perceive the risk of public listing (as Berlingske went through) as one worth running, because it also means cheap cleaning. Hence, a challenge in the collaborative or interactive regulation is that the consumers can behave unexpectedly or simply ignore the 'nudge' (Thaler & Sunstein, 2008) and 'motivation' to act. This makes the regulation *less regular* and more flexible or random, because some purchasers may act but others may not. This is a challenge in a society where we cherish equality to the law, and where we expect systematic regulation and enforcement from tax administrations. Enforcement through the taxpayers and third parties may introduce an element of regulatory randomness.

PUBLIC LISTING AS A TOOL FOR PRESSURE

As described previously, the case workers function less as classic auditors and more as meta-governors when they provide consumers with information that prompts them to act. Some of the consumers/purchasers take action voluntarily in this set-up, such as when the purchasers receive a guidance visit or read the stories about non-compliance contracting in the media. These purchasers may revise and check up on their own contracts and cleaners, because they see that there is a problem and they do not wish to be exposed in the media if they were suddenly targeted in a random audit. What unites these consumers' actions, and what is strategically advocated by the Tax Administration, is that these consumers' care for their public reputation. The Tax Administration connects to the purchasers' care for CSR.

Although this seems to be a legitimate way of motivating or 'nudging' consumers to take on responsibility, other purchasers are pressed to take action in a more problematic fashion. Berlingske is an example of this, because this organisation was involuntarily named in the media. Significantly, the name of 'Berlingske' was not exposed by the tax administration. However, by publicly announcing that an inspection had taken place in a *newspaper group* within the Danish *media industry*, the Tax Administration narrowed the group of potential organisations. In fact, there are perhaps just two or three newspaper groups in Denmark; thus, figuring out which one of these recently experienced a 'tax raid' should not be a problem for a curious journalist. What is problematic is that the purchasers of cleaning services are promised anonymity in the press releases. Yet, based on the anonymised information, it is possible to list the parties involved. In this way, it is not the Tax Administration that does the listing, but the Tax Administration plants the seeds for the public disclosure.

There are studies that have investigated the functioning of public listings. Researching public tax blacklists, Sharman notes that this is a form of 'speech act' that changes the world by "damaging states' reputations among investors, and this produced pressure to comply through actual or anticipated capital flight" (2009: 573). Sharman notes that public blacklisting can be an effective means of bringing about compliance in otherwise recalcitrant states. Yet, there are also challenges connected to listings. Sharman and Rawlings show that listings are often arbitrary and discriminating, meaning that blacklists are not compiled objectively and those listed are not selected based on a consistent set of rules or criteria (Sharman & Rawlings, 2006). These findings are interesting to connect to this study. Significantly, this study does not concern direct blacklisting of specific jurisdictions, taxpayers or corporations - such as, for instance, the national tax blacklists of Guernsey, Cayman or the Isle of Man (e.g. Sharman & Rawlings, 2006). Yet, the snowballing effect that comes from the unannounced inspection at Berlingske is that this organisation's name became listed publicly. In line with Sherman and Rawling's study, this case highlights that this listing is an effective means of bringing about compliance, as Berlingske changed its supplier of cleaning services because it feared reputational damage.

The challenge of using public listing as a tool for 'pressure' in this case is that the public listing is slightly off target, because the listed consumer has not done anything wrong. Yet, this consumer is enrolled in the Tax Administration's regulatory craft precisely because this consumer's perception is that its name should not be connected to tax evasion and the human exploitation of cleaners. Relating the case further to the above research, it is also a point that the listing in this case is arbitrary. The Tax Administration produces its press releases, but it does not know which

of the incidents will develop into media scandals because this is dependent on others' actions - for instance, independent journalists. Hence, as noted in the previous section, the interactive regulation may be problematic because it lacks regularity and consistent application, and because it encourages the listing of guiltless consumers. If anyone should be listed publicly, it ought to be the non-compliant cleaning companies and sub-contractors who are responsible for the tax evasion. Yet, as these have little reputational care and CSR awareness, there is not much potential in listing them. The potential for the Tax Administration lies getting the sensitive consumers listed.

MOVING COMPLIANCE COSTS TO THE EXTERNAL PARTNERS

In a discussion of user involvement, co-production and collaboration, Ryan notes that such initiatives are often initiated when other forms of service delivery have not achieved the goals and objectives in the relevant area of policy (2012, p. 319). This is also the starting point for the case presented here: what was done prior to enforcing the legal code turned out to be ineffective and a new mode of regulation was called upon instead. A central element of the new regulation was that it shifted part of the 'burden' of the regulation to the consumers. Hence, as Ryan notes, co-production or collaboration mobilises consumers, because this reduces the burden on public resources. In the case reported here, collaboration with the consumers is not simply initiated because there is an obligation to engage consumers and the citizenry to participate in the policy process, but these groups are, to a large extent, engaged because they take on regulatory responsibilities. The pre-emptive actions of Berlingske and all of the other consumers reduce the costs of tax compliance for the Tax Administration because there is less need for the resource-demanding reactive audits when the consumers themselves proactively take steps to ensure tax compliance. A central point, and reason why the tax inspectors find the approach compelling, is that it supplements and aids the Tax Administration's own regulation efforts.

Although this might be convenient for the Tax Administration, Slemrod draws attention to the challenges that arise when 'administrative costs', such as the costs endured by the Tax Administration to ensure tax compliance, are pushed over to become taxpayers' or third parties' 'compliance costs' (2015). This is what occurred in this case, and a caveat to this strategy is that it might make the Tax Administration look more efficient and less costly, but it does not necessarily lower the total cost of ensuring tax compliance. Instead, the burden of the regulation is simply delegated to the purchasers, who now need to spend resources on checking contracts, IDs and salary payments. As Slemrod notes, this is problematic because it is difficult to measure the compliance costs of the taxpayers and third parties (ibid. 12). A challenge with the collaborative regulation strategy is thus that it delegates or outsources compliance costs to a 'partner' who, in many instances, cannot refuse to take on the responsibility and who may not be able to document the amount of resources it uses on the regulatory efforts.

THE POTENTIALLY WEAK REVENUE FOCUS

The last challenge with collaborative or interactive regulation is that it has a potentially weak relationship to the traditional revenue-raising function of the Tax Administration. In a recent article in the *Journal of Tax Administration*, Hickman notes that the Inland Revenue Service (IRS) administers government programmes that have little to do with the traditional raising of revenue and more to do with programmes, purposes and functions of social welfare (2015). She explains that the IRS has become one of the U.S. government's principal welfare agencies because it tries

to ‘alleviate poverty’ and ‘support working families’, and ‘subsidizes approved activities’. Although these are important challenges in today’s society, Hickman notes that it is worth noticing that the IRS is focused on pursuing goals and administering programs with only a tangential relationship to the U.S. tax system’s traditional revenue-raising mission.

The purpose of the regulation of contracting in the cleaning sector is to *restrict* the market for ‘shady’ sub-contractors of cleaning services by influencing the consumers to take responsibility for their cleaning and pay closer attention to their contracts. The project description notes that this ‘change of attitude’ in the consumers is believed to help minimise the tax gap. The tax gap is indeed connected to the traditional revenue-raising function of the Danish Customs and Tax Administration, yet neither the project description nor the protocol for the work contain any targets or aims connected directly to revenue collection. All revenue collection is perceived to be attained indirectly when tax-evading cleaning companies and subcontractors close and tax-compliant cleaning companies are contracted with instead. Furthermore, importantly and closely related to Hickman’s point, the discourse about this regulatory work is focussed on how it will prevent exploitation of cleaners and how it will change the market for contracting of cleaning services. This indeed resembles Hickman’s point that the “IRS has transitioned over time from a mission-driven agency that collects taxes to an omnibus agency that does many things” (ibid. 74). She fears that the IRS may reach an organisational tipping point where the agency’s resources are being stretched too thinly between too many goals. Relating this to the Danish case, it is significant that the work emphasises that the regulation also stops the exploitation of foreign workers. This more resembles the ‘social welfare kind’ of work than direct revenue collection. This indicates that the Danish Tax Administration also acts as a welfare agency: it hopes that its activities will hinder human exploitation, and in this way, the agency focusses on solving broader social challenges in society by means of its regulation. Whether such an ‘omnibus agency’ is to be desired or not is up to the future to decide.

CONCLUSION

In this paper, I have shown that the Danish Customs and Tax Administration collaborates with consumers in its regulation of non-compliance in the cleaning sector. My main argument is that this is an instance of collaborative and interactive governance; a form of governance that is widespread in other domains of public administration but not often seen or used within tax administration.

My argumentation has several steps. First, I show that the regulation relying on direct and reactive enforcement targeting the tax-evading subcontractors was deemed ineffective by the case workers. The reason was that when these subcontractors were charged by the Tax Administration, they either fled the country, had no assets with which to pay their tax debts, or simply closed their business to start a new ones in other names. Hence, the direct and reactive regulation was not able to stop the activities. Second, I show that a new regulatory set-up has been embarked on. This is a set-up that targets the consumers of cleaning services. These consumers or purchasers of cleaning services are *not* formally or legally responsible for the tax compliance of either the cleaning company they sign their cleaning contract with or the hired subcontractors who provide the cleaners. Yet, the Tax Administration’s idea is that *if* an incentive could be installed in these consumers whereby they would start to take responsibility for the tax compliance of their cleaning

companies and their subcontractors, *then* much would be gained in preventing tax evasion from occurring in the first place.

Thirdly, I show that the Tax Administration uses two tools to motivate the consumers to take responsibility. The first tool is a number of guidance meetings, in which the inspectors provide information about the challenges in the cleaning sector and what the purchasers can do to prevent the human exploitation of cleaners and tax evasion. These meetings are only briefly described in the analysis. The second tool is a number of unannounced inspections and subsequent press releases. The functioning of this tool is described using the case of Berlingske. Here, we see that a respectable and CSR-sensitive purchaser of cleaning services (unwillingly) becomes listed publicly. This consumer becomes connected to stories of police arrest, illegal workers, exploitation and tax evasion. This motivates (or forces) Berlingske to take action and change to a compliant supplier of cleaning services because it fears reputational damage from not reacting.

What these two tools do together is to motivate the consumers to take responsibility for the cleaning that happens at their property - despite the fact that they have no legal responsibility. Roughly speaking, the guidance meetings can be characterised as a 'carrot' (a positive pull to act pre-emptively) and the unannounced inspections and press releases as a 'stick' (that is, as a deterrent push to act because none of the consumers want to be publicly listed, and connected to tax evasion and human exploitation). That the public listing is *not done* by the tax administration, but by independent journalists, does not change the situation for the listed consumers. The consumers are in an unfortunate situation in any case. This way of using a stick to achieve tax compliance follows Sharman's research on tax blacklisting. He notes that blacklisting is not just cheap talk of signalling, but a stick that can be used to force non-compliant actors into compliance (Sharman, 2009, p.593). I see the same effect here, except that the blacklisting and the 'stick' are constructed interactively, as it is the independent journalist that does the listing based on the Tax Administration's anonymous and unannounced inspections. To sum up, in their different ways, these tools get the otherwise guiltless consumers to act by pre-emptively checking up on their contracting of cleaning services.

The aspiration of the Tax Administration is that the outcome will be (or is) that the non-compliant cleaning companies and the subcontractors' cash flows are reduced, leading these businesses to either close or comply. The Danish Customs and Tax Administration perceives this regulation as efficient because it draws in the (regulatory) resources of a number of external partners who assist the Tax Administration in regulating an area that has turned out to be problematic, and which is ripe for human exploitation and tax evasion.

In the section "Interactive governance - and its challenges", I argue that the new regulatory set-up is an example of collaborative governance, because the regulation is done interactively with the consumers. The case workers function as meta-governors who provide information that nudges the consumers to act. In this way, there is (ideally) no direct contact between the Tax Administration and the tax-evading subcontractors. Instead, the tax evaders are 'played out' by the actions of the consumers (for instance, when the consumers insist that salaries be paid to bank accounts or when they start to check for false IDs). In this set-up, the consumers are 'co-producers' or 'partners' in the regulatory craft. What this new regulatory set-up aptly illustrates is that a plurality of actors

(state actors together with non-state actors) interact to achieve the common objective of tax compliance. This, indeed, is the hallmark of interactive governance.

CONTRIBUTION

The contribution of the article is twofold. First it shows an instance of *collaborative* and *interactive governance* within tax administration. This is interesting to analyse, as some scholars express that *participation, collaboration, partnerships* or *interaction* is something that happens in the ‘servicing’ part of the public sector (school, hospitals or city planning) but not in the ‘coercive’ parts, such as in the police, military or tax administration. Recall how Sparrow writes that the role of a tax administration is to impose duties and to deliver obligations, rather than to deliver services, or Torfing, who states that interactive governance is not appropriate for all policy areas, especially not law, defence and taxation. This article’s contribution to this general discussion about collaborative and interactive governance is to show that this regulation exists within tax administration, and that this new form of governance does not prevent either the imposition of duties or the fulfilment of obligations. The fact that the methodology for enforcing the regulations has changed does not mean that any of the legal obligations of the cleaning companies or subcontractors have changed. They are still responsible for their own tax compliance.

The second contribution of the article is that it opens *collaborative* and *interactive governance* to criticism. It does so by showing four challenges related to the regulation. First, the regulation may result in regulatory randomness, as the tax administration cannot control which consumers act or how they act. Secondly, the analysis problematizes that some consumers (such as Berlingske) end up being publicly listed despite being guiltless. This *indirect* public listing seems to be different from normal tax blacklisting, where it is the non-compliant actors who are listed. Third, the article raises the concern that the costs/resources used for the regulation are simply shifted from the tax administration to the consumers. This makes the tax administration look more effective - it achieves more with less. However, this is a ‘false’ equation because the endured costs are placed with the consumers, who may have difficulties documenting the amount of resources they use in checking up on contracts, and keeping an eye on their cleaners’ IDs and papers. Finally, I use Hickman’s point about how tax administrations focus less on the traditional raising of revenue and more on programs, purposes and functions that have to do with social welfare. I argue, that there is a similar movement in the Danish case, because the regulation targets a mix of challenges related to moonlighting, tax evasion and the human exploitation of vulnerable cleaners.

These points of criticism are relevant to this case, yet my hope is that this case and its challenges can also contribute to a wider discussion and problematization of the use of *collaborative* and *interactive governance* in other domains of the public sector. Finally, recall this article’s subtitle: "Can consumers be engaged as ‘partners’ in the regulatory craft?" The answer to this question is that consumers can be (and are) engaged as partners and collaborators in interactive regulation. Yet, such engagement must be evaluated and monitored continually, as there is a problematic ‘flipside’ to this new and innovative form of governance.

REFERENCES

- Alford, J. (2009). *Engaging public sector clients: from service-delivery to co-production*. Basingstoke, Hampshire: Palgrave Macmillan Ltd.
- Alford, J. (2013). The Multiple Facets of Co-Production: Building on the work of Elinor Ostrom. *Public Management Review*, 16(3), 299-316.
- Boll, K. (2014). Shady car dealings and taxing work practices: An ethnography of a tax audit process. *Accounting, Organizations and Society*, 39(1), 1-19.
- Boll, K. (2015). Deciding on tax evasion – front line discretion and constraints. *Journal of Organizational Ethnography*, 4(2), 193-207.
- Boll, K. (2016). Decentring regulation. Tax fraud and the cleaning sector. In R. A. W. Rhodes & M. Bevir (Eds.), *Rethinking Governance: Ruling, Rationalities and Resistance* (pp. 179-197). London, United Kingdom: Routledge.
- Boll, K., & Rhodes, R. A. W. (2015). Excursions in administrative ethnography. *Journal of Organizational Ethnography*, 4(2).
- Boll, K., & Tell, M. (2016). Proactive Public Disclosure: A new regulatory strategy for creating tax compliance? *Nordic Tax Journal*, 2015(2):36-46 DOI [10.1515/ntaxj-2015-0009](https://doi.org/10.1515/ntaxj-2015-0009).
- Bovaird, T. (2007). Beyond Engagement and Participation: User and Community Coproduction of Public Services. *Public Administration Review*, 67(5), 846-860.
- Dunston, R., Lee, A., Boud, D., Brodie, P., & Chiarella, M. (2009). Co-Production and Health System Reform – From Re-Imagining To Re-Making. *Australian Journal of Public Administration*, 68(1), 39-52.
- Fledderus, J., Brandsen, T., & Honingh, M. (2013). Restoring Trust Through the Co-Production of Public Services: A theoretical elaboration. *Public Management Review*, 16(3), 424-443.
- Grabosky, P. N. (1995). Using Non-Governmental Resources to Foster Regulatory Compliance. *Governance*, 8(4), 527-550.
- Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in Practice*. London, New York: Routledge.
- Harrits, G. S., & Møller, M. Ø. (2013). Prevention at the Front Line: How home nurses, pedagogues, and teachers transform public worry into decisions on special efforts. *Public Management Review*, 16(4), 447-480.
- Hickman, K. E. (2015). Administering the Tax System We Have. *Journal of Tax Administration*, 1(1).
- Hood, C. (2005). Public management. The word, the movement, the science. In E. Ferlie, L. E. Lynn, & C. Pollitt (Eds.), *The Oxford handbook of public management* (pp. 7-26). Oxford; New York: Oxford University Press.
- Jakobsen, M., & Andersen, S. C. (2013). Coproduction and Equity in Public Service Delivery. *Public Administration Review*, 73(5), 704-713.
- Lipsky, M. (2010 [1980]). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.

- Moeran, B. (2005). *The business of ethnography: strategic exchanges, people, and organizations* (English ed.). Oxford; New York: Berg.
- Needham, C. (2008). Realising the Potential of Co-production: Negotiating Improvements in Public Services. *Social Policy and Society*, 7(02), 221-231.
- Neyland, D. (2007). *Organizational Ethnography*. Thousand Oaks, California: Sage Publications.
- Oats, L. (Ed.) (2012). *Taxation: A Fieldwork Research Handbook*. London and New York: Routledge.
- OECD. (2013). *Together for better outcomes: Engaging and Involving SME Taxpayers and Stakeholder*. Retrieved from Forum on Tax Administration Compliance Sub-Group.
- Podger, A. (2012). Putting Citizens First: A Priority That Needs to be Addressed with Care. *Australian Journal of Public Administration*, 71(1), 85-90.
- Rhodes, R., A. W. (1997). *Understanding governance: policy networks, governance, reflexivity, and accountability*. Buckingham; Philadelphia: Open University Press.
- Rhodes, R., A. W. (2011). *Everyday life in British government*. Oxford, New York: Oxford University Press.
- Rhodes, R., A. W. (2015). Ethnography. In M. Bevir & R. A. W. Rhodes (Eds.), *Routledge Handbook of Interpretive Political Science* (pp. 171-185). Abingdon, Oxon: Routledge.
- Ryan, B. (2012). Co-production: Option or Obligation? *Australian Journal of Public Administration*, 71(3), 314-324.
- Sharman, J. (2009). The bark is the bite: International organizations and blacklisting. *Review of International Political Economy*, 16(4), 573-596.
- Sharman, J., & Rawlings, G. (2006). National Tax: A Comparative Analysis. *The Journal of International Taxation*, 17(9), 38-47.
- SKAT. (2011). *Rengøringsmedarbejdere i bladhus blev anholdt*. Copenhagen: SKAT.
- Slemrod, J. (2015). Sexing Up Tax Administration. *Journal of Tax Administration*, 1(1).
- Sparrow, M. K. (2000). *The regulatory craft: controlling risks, solving problems, and managing compliance*. Washington, D.C.: Brookings Institution Press.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*: Penguin Books Limited.
- Torring, J., Peters, B. G., Pierre, J., & Srensen, E. (2012). *Interactive Governance: Advancing the Paradigm*. Oxford: Oxford University Press.
- Unger, B. (2009). Money Laundering – A Newly Emerging Topic on the International Agenda. *Review of Law & Economics*, 5(2), Article 1.
- Unger, B. (2013). Can Money Laundering Decrease? *Public Finance Review*.
- van Eijk, C. J. A., & Steen, T. P. S. (2013). Why People Co-Produce: Analysing citizens' perceptions on co-planning engagement in health care services. *Public Management Review*, 16(3), 358-382.

- Yanow, D., Ybema, S., & van Hulst, M. (2012). Practicing Organizational Ethnography. In G. Symon & C. Cassell (Eds.), *Qualitative organizational research, Core methods and current challenges*. London: Sage.
- Ybema, S., Yanow, D., & Kamsteeg, F. (Eds.). (2009). *Organizational Ethnography. Studying the Complexities of Everyday Life*. Los Angeles: Sage.

EMPIRICAL EVIDENCE ON COOPERATION BETWEEN SUB-CENTRAL TAX ADMINISTRATIONS

José María Durán-Cabré, Alejandro Esteller-Moré¹ and Luca Salvadori²

Abstract

The literature on horizontal tax interdependence pays limited attention to interactions in administrative policies, although they can play a large role in determining the amount of tax revenues collected. We investigate the incentives for sub-central tax authority cooperation in a decentralised context, with the aim of identifying the determinants of that cooperation. Our results are congruent with standard theory; in particular, the existence of reciprocity is essential for sharing tax information, but there is sluggishness in this process, which is partly the result of the short-sighted behaviour of tax authorities influenced by budget constraints. Hence, this is good news for the functioning of a decentralised tax administration as, in the medium to long run, the gains to be made from sharing tax information are achieved.³

INTRODUCTION

Tax administration policies are crucial in determining the final amount of revenues collected by tax authorities. Furthermore, be it in a federal context with decentralised tax administrations, or internationally with different national administrations, tax authorities are dependent on each other to enforce tax rules. For example, starting from January 2017, all EU member states will have to automatically exchange information on tax rulings given to companies with cross-border operations. The aim is to provide national authorities with insight into aggressive tax planning in order to protect their tax bases; consequently, cooperation will be essential. In general, improving tax enforcement in the global economy has translated into a proliferation of bilateral and multilateral treaties between national tax administrations and tax havens. Given these circumstances, investigating the determinants of cooperation between tax administrations has

¹ Universitat de Barcelona & Institut d'Economia de Barcelona (IEB).

² Universitat Autònoma de Barcelona, IEB & Tax Administration Research Centre (TARC). Corresponding author: Department d'Economia i d'Història Econòmica, Unitat de Fonaments de l'Anàlisi Econòmica (UFAE), Edifici B – Campus Universitari s/n, 08193 Bellaterra (Barcelona), Spain; e-mail: luca.salvadori@uab.cat.

³ ACKNOWLEDGMENTS. We would like to thank participants at the ZEW 2014 Public Finance Conference, at the “XXI Encuentro de Economía Pública” (University of Girona), at the ERSA 2013 congress, at the ERSA 2012 Summer School, at the SIEP 2012 conference, and at the PhD seminars of the IEB, CREIP (University of Tarragona) and Max Planck Institute for Tax Law and Public Finance (Munich) for their helpful comments. We are also very grateful to Nadja Dwenger, Salmal Qari, Federico Revelli, Lionel Védrine and two anonymous reviewers for their useful suggestions. The funding from the Generalitat de Catalunya (2009SGR102) and from the Ministerio de Ciencia e Innovación (ECO2012-37873) are gratefully acknowledged. Durán-Cabré also thanks the funding from the Institut d'Estudis Autònoms. Usual disclaimers apply.

become a key issue; yet, the literature on horizontal tax interdependencies pays limited attention to these matters.⁴

We seek to investigate the potential for cooperation in tax administration policies between sub-central tax authorities by carrying out an empirical analysis in a federal context. This represents something of a novelty in the literature and should serve to shed some light on alternative designs (centralised vs. decentralised) for tax administration within this context. In doing so, we analyse the determinants of information sharing between regional administrations based on the Spanish case, which is a good field for empirical research. Spanish regions (the so-called “Comunidades Autónomas”, henceforth CAs) have had the power to administer several wealth taxes⁵ since the mid-1980s and, following reforms in 1997 and 2002, have also acquired the legislative power to modify significant statutory tax parameters.⁶ Thus, this case study should serve as a benchmark for evaluating the information-sharing process in a decentralised framework and, more generally, for analysing the efficiency of a decentralised tax administration scheme.

We focus our empirical analysis on a specific area of potential cooperation between the CAs, the only one for which official data is available. In the case of wealth taxation, legal tax allocation principles (in Spanish, the so-called “puntos de conexión”) indicate how tax revenues should be distributed among the CAs: the residence principle and the territorial (or source) principle, depending on the taxable event.⁷ However, taxpayers are not necessarily aware of these and so might commit errors when reporting their tax returns: that is, a taxpayer might pay the tax to the wrong CA.⁸ Thus, each CA should share their information on misreported taxes and transfer the corresponding revenue to the competent CA. This is supposedly an automatic practice, but in reality it does not always occur this way. Indeed, there is considerable casual evidence confirming that the information sharing process between CAs is far from automatic.⁹

⁴ The effectiveness of these cooperative policies at an international level has been highly questioned by recent empirical literature (e.g. Johannesen & Zucman, 2014). As we will make clear though, our approach is different, as we focus on the administrative incentives to cooperate within a federation.

⁵ Namely the inheritance and gift tax (IGT), the annual wealth tax (AWT) and the tax on wealth transfers (TWT).

⁶ For more details on these reforms, see Esteller-Moré (2008).

⁷ In the case of the IGT, three different circumstances may occur. The residence principle applies to all inheritances: the tax revenues are collected in the CA of residence of the deceased. This principle also applies for gifts of chattels but the relevant residence in this case is that of the donor. Finally, in the case of the gift of real estate, the territorial principle applies. The AWT is based on the residence principle, while the TWT is mainly based on the territorial principle.

⁸ Suppose, for example, that a company with its headquarters in Madrid sells a block of flats located in the CA of Andalusia and pays the TWT to the CA of Madrid. In this case, an error has occurred, as the TWT is subject to the territorial principle and the tax return should be reported to the CA of Andalusia. Similarly, there is a mistake when a daughter living in the CA of Valencia receives an inheritance from her father, whose residence was in the CA of Catalonia, and she reports the IGT to the region in which she lives, rather than to Catalonia, as she should have according to the allocation principle.

⁹ Every year, tax inspectors from the State review the way in which each region administers its ceded taxes and they report their findings in the “Informe sobre la cesión de tributos a las Comunidades Autónomas”. For instance, in the 2006 report about Catalonia, inspectors from the State explain: “It should be noted that existing experiences show an unequal behaviour of the different CAs in their degree of compliance with the obligation to submit the information and the income due to the competent CA. The perception that the competent services of the Directorate General of Taxes of the Catalan government have on this issue is that certain CAs systematically and, in many cases, violate that obligation” (p.39 of the report). Moreover, from informal conversations maintained with former directors of the Catalan Tax Authority, we know that in some cases they chose not to transmit information to other CAs until the latter

This situation might arise because every CA faces a trade-off between, on the one hand, cooperating by transmitting the information and the misreported tax revenues to other CAs, and, on the other, not cooperating and retaining the misreported tax revenues. The costs of cooperation are mainly administrative (being related directly to this information-sharing process) and financial (a loss of revenue yields). The benefits of cooperation are based on reciprocity: if a CA cooperates, it might foster other regions' cooperation in the future. For this reason, if a CA does not cooperate, there may be a cost, as the other CAs will opt not to exchange information in the future. In a repeated game, cooperative behaviour should produce mutual benefits for both CAs, since the benefits due to reciprocity should be higher than the administrative and financial costs in the short-run. Therefore, our main hypothesis is that a CA's cooperative behaviour is a matter of reciprocity, as it depends strictly on the potential cooperation of the other CAs in previous periods.

To test this hypothesis, we estimate a Tobit random-effect model and also a dynamic version of this model to account for sluggish adjustment in transmitted tax revenues. Our results confirm the role played by reciprocity and indicate the presence of persistency in the strategic behaviour of the tax administration. In addition, in keeping with the short-run financial benefits of non-cooperation, we find that the impact of reciprocity is lower when the CAs face budget constraints picked up by the deficit. Thus, according to our analysis, in the medium to long run, the regional administrations learn the advantages of cooperation, thus providing elements that support the correct functioning of a decentralised tax administration.

The rest of the paper is organised as follows: section 2 provides a summary of the relevant literature; in section 3 we present our empirical strategy; section 4 presents the results; and we conclude in section 5.

LITERATURE REVIEW

The literature has identified two main sources of horizontal interdependence at a tax administration level.¹⁰ On the one hand, Cremer and Gahvari (2000), examining the implications of tax evasion for fiscal competition and tax harmonisation policies in an economic union, demonstrate the possibility of mobility-based competition in tax enforcement policies. They obtain sub-optimal equilibrium values for both tax and audit rates, and show that tax harmonisation alone is not sufficient to avoid strategic incentives to attract tax bases, as there can be no commitment to audit policies. Durán-Cabré, Esteller-Moré and Salvadori (2015) have tested this result for the Spanish decentralised framework and corroborate the presence of mobility-based competition in tax enforcement among regional administrations.

On the other hand, the incentive for sub-central tax authorities to collaborate by sharing relevant tax information has also been accounted for in the literature that has focussed on the incentives for tax cooperation between countries to reduce evasion in an international mobile-capital framework

opted to do the same with their misreported taxes. This seems to suggest that 'reciprocity' might play a relevant role in determining the extent to which information is shared between CAs. Indeed, in the 2002 report about another CA, Castile and León, the inspectors from the State explain that this region would not return revenues due to the CA of Madrid until the latter transferred revenues due to it.

¹⁰ More generally, recent literature has also identified the incentives for vertical transmission of information between central and local governments in a federal framework. Dreher, Gehring, Kotsogiannis and Marchesi (2014) explore the role of this information transmission process in explaining the optimal degree of decentralisation across countries.

(see Keen & Ligthart (2006a) for a survey). In particular, the seminal study by Bacchetta and Espinosa (1995) identifies the strategic trade-off between competitive behaviour (lowering the tax rate to increase foreign investment) and cooperative behaviour (voluntarily sharing information to reduce international tax evasion). In equilibrium, the second effect may dominate the former, resulting in partial information exchange. In a more recent study, Bacchetta and Espinosa (2000) further their previous analysis by modelling the choice of tax rates and information provision as an infinitely repeated game. A contribution in this same line is provided by Huizinga and Nielsen (2002), who model a repeated game in which tax authorities choose between withholding taxes and sharing information as alternatives for dealing with international capital income and profit taxation.¹¹

Both studies argue that potential cooperation in information sharing is a matter of reciprocity and, in particular, that it may be sustained if the process is viewed as an infinitely repeated game rather than as a single one. In this regard, the propensity of a country to cooperate directly depends on the potential cooperative behaviour of the other country in previous periods. Thus, in these models, each country evaluates the trade-off between not providing information and obtaining a corresponding temporary gain (due to their attracting tax evading investors) versus suffering the costs of the non-cooperative behaviour of the other country (generally, more aggressive tax competition, or the absence of information exchange, or both) forever after.

Our empirical framework reflects existing theoretical models – given the existence of a trade-off between cooperative and non-cooperative behaviour – but applied to a federal context. The main differences between the two contexts lie in the tax authorities' motivation and incentive to cooperate. In an international framework with mobile capital, countries share fiscal information with the aim of avoiding, or of at least reducing, a race to the bottom in tax rates and the resulting negative effects on tax revenues. This kind of cooperation between countries reduces tax fraud. In our federal context, we focus on the potential existence of cooperation in tax administration between sub-central authorities. This is probably more related to tax-management policies than to strategic behaviours of the regional tax authorities.¹² In a federal framework, a decentralised tax administration might enhance efficiency due to a greater ability of sub-central authorities to exploit informational advantages on local tax bases (see, for example, Martinez-Vazquez & Timofeev, 2010). Nonetheless, this is conditional upon the existence of cooperation among sub-central tax authorities. That is why it is so important to test for the existence of administrative cooperation.

Some empirical papers have tested these models in an international framework. In particular, Ligthart and Voget (2010) study the determinants of tax information sharing between Dutch and foreign tax authorities for income tax purposes. From our perspective, the most interesting result in this paper concerns reciprocity. The authors show that an increase in the amount of tax information provided by the Dutch tax authorities to their foreign counterparts significantly increases the amount of information received by the Dutch tax authorities. Elsayyad (2012) analyses recent treaty signings between tax havens and OECD countries as the outcome of a bargaining process over treaty form and focusses on the presence of an exchange of information

¹¹ These contributions generated further research (e.g. Tanzi & Zee (2001); Chisik & Davies (2004); Keen & Ligthart, (2006b).

¹² Note that cooperation by sharing information on misreported taxes should not have implications concerning taxpayers' compliance.

clause. The paper shows that the likelihood of treaty-signing is mainly driven by a tax haven's bargaining power and good governance. Moreover, the author finds that it is easier for an OECD country to renegotiate an existing treaty so as to incorporate an information exchange clause than to pressure countries to do so without an existing agreement. By interpreting the existence of a previous agreement between two countries as a measure of reciprocity, we have further confirmation that reciprocity matters in determining the level of information exchanged between two tax authorities.

In our federal framework, CAs are required to cooperate by law in support to the good organisation of the federal tax administration system. To this aim, sub-central tax authorities should automatically rectify any errors that might arise in the reporting of tax returns, but they have an incentive not to cooperate due to the presence of administrative costs and to the loss of financial revenue yields. In this context, and according to our hypothesis, reciprocity should reinforce the tax information exchange process by being an important driving force in the promotion of cooperation and the enhancement of the functioning of the decentralised tax administration. This empirical analysis of a federal framework represents, we believe, a novelty and progress in the literature.

EMPIRICAL ANALYSIS

In this section, we present the dataset and define the empirical methodology employed in developing our analysis.

The Empirical Framework

Data on Spain's regional tax administrations are extracted from the report "Informe sobre la cesión de tributos a las Comunidades Autónomas", published every year jointly with the project of the general State budget. Specifically, we have access to data on the total number and total amount of transfers resulting from misreported tax returns ("Transferencias por aplicación de los puntos de conexión") collected (returned) by each CA from (to) any other region during the 1989-2009 period.¹³ Hence, in contrast with previous analyses, our dataset allows us to identify both directions in the information-sharing process. Additionally, the availability of a time span allows us to adopt a dynamic approach and, thus, to test for the possibility that regional administrations learn the potential advantages of gradually sharing information.

Our endogenous variable is the amount of tax revenues transferred by each CA to every other CA in a given year and thus takes the form of a continuous random variable over strictly positive values, but it assumes the value zero with positive probability. Our dataset contains 43.02 percent zero-valued output. Thus, our endogenous variable may be censored at zero inasmuch as a zero value could alternatively indicate an actual absence of misreported taxes, or that CAs choose not to share information on misreported taxes and claim to have zero tax revenues to transmit.

¹³ For instance, in 2000, the region of Andalusia transferred 828,192 euros to the region of Castile-La Mancha, corresponding to seven cases of misreported taxes. And the latter, for example, transferred 15,872.9 euros to the region of Valencia, corresponding to 33 cases.

Therefore, we maintain the random-effects Tobit corner-solution model as our main approach (see Wooldridge, 2002, pp. 518-549)¹⁴, which is defined as follows:¹⁵

$$Trans_Rev_{ijt} = \max[0, \alpha Rec_Rev_{ijt-1} + Y_{ijt}\beta + X_{it}\mu + \tau_t + \vartheta_{ij} + \varepsilon_{ijt}] \quad (1)$$

where $Trans_Rev_{ijt}$ is the amount of misreported tax revenues transmitted by region i to region j during year t . We control for reciprocity through the misreported tax revenues received by region i from region j during the previous year, Rec_Rev_{ijt-1} . This is the key regressor, since our main hypothesis is that reciprocity fosters cooperation between regional tax authorities and then we expect α to be positive. More precisely, if α is equal to zero, receiving tax revenues from CA j does not encourage a transmission of revenues by CA i . This could mean that CA i does not have any case of misreported taxes to transmit or that it does not have incentives to do it. Instead, a strictly positive α surely indicates that there are cases of misreported taxes and, most importantly, that CA i has some incentives to transmit the corresponding misreported tax revenues.

We introduce a series of control variables that account for both region pair-specific characteristics and unilateral determinants referring to region i that might influence the information-sharing process. The pair-specific variables are collected in vector Y_{ijt} . In particular, N_{ijt} is the number of cases of misreported taxes transmitted from region i to region j in year t . According to Ligthart and Voget (2010), the distance between regions might reduce the flow of information between them. Indeed, this variable accounts for both higher transaction costs and lower cultural proximity, which are relevant issues in a federal framework.¹⁶ Therefore, we control for D_{ij} the physical distance in kilometres between i and j . The political alignment between Spanish regions¹⁷ is another variable that might have an impact on the tax administrations' willingness to cooperate. Thus, we introduce PA_{ijt} , a dummy identifying the political alignment between the two regions at time t . The relative GDP of the two regions at time t , $RGDP_{ijt}$ is also included in order to account for the relative economic power of the two regions; that is, as a measure of the relative bargaining position of region i with respect to region j (Elsayyad, 2012). A positive (negative) sign would indicate a favourable (unfavourable) bargaining position of region i with respect to region j due to a higher (lower) amount of revenues transmitted by region i to region j .

The vector X_{it} includes a constant term and the unilateral variables. According to the previous literature on the exchange of tax information (Bacchetta & Espinosa, 1995, 2000), the statutory tax parameters and the enforcement costs are crucial in determining the level of information

¹⁴ In a previous version of this paper, we employed the number of cases of misreported taxes transmitted as our endogenous variable. Given that this is a count-data variable, we used an estimation strategy based on Poisson regression models obtaining results that are congruent with those obtained through the current estimation strategy. These results are available upon request.

¹⁵ A limit of our database is intrinsic in the absence of a counterfactual: we are not able to disentangle a priori if a zero valued number of cases of misreported taxes is due to an actual absence of cases of misreported taxes or to a strategic uncooperative behaviour. Nevertheless, the methodology we employ appropriately takes this shortcoming into account.

¹⁶ On the one hand, since we analyse a long period of time, including many years before the "internet era", this variable is relevant in measuring larger operational costs due to a longer distance between two regions. On the other hand, the distance between regions in the same federal country might also be relevant in representing the level of cultural affinity.

¹⁷ Please note: this factor is specific for an analysis within a federal context.

exchange between tax authorities. These issues are also relevant in our context, albeit in a different way; thus, we control for $Tot_Reg_Tax_Revenues_{it}$ and $Tot_Reg_Audit_Revenues_{it}$ that account for total tax revenues and total tax auditing revenues collected by region i during year t , respectively. These variables are proxies of regional tax autonomy in raising revenues and they are expected to be associated with greater amounts of information being exchanged. Budgetary and political variables might also play a role in determining tax administration policies (see, for example, Esteller-Moré, 2005, 2011). In particular, we control for the deficit expected at the beginning of every fiscal period in order to account for the financial conditions of regional budgets and to measure indirectly the financial opportunity cost of cooperation of region i . We expect a higher deficit to negatively impact the transmission of misreported revenues. We return to this variable below. We include the total amount of transfers received from the central government divided by total regional expenditure to account for a further budgetary factor relevant in a federal framework, such as that operated in Spain. We expect this variable to have an income effect on the behaviour of the tax administrations. In particular, a higher transfer-expenditure ratio should force the administration to rely less on its own tax resources and to transfer more tax revenues to the other regions.

We are not able to identify the impact of the administrative costs of cooperation, but reasonably suppose it to be constant over time. As such, it will be picked up by the constant term; however, if it varies over time (and uniformly throughout the CAs) it will be picked up by the time effects. In the case of the political variables, we include a dummy equal to one, El_{it} , if there is a regional election in i CA during the year t , to control for the potential impact of the electoral cycle on the incentives to share information. To account for modifications to the statutory tax parameters, we include a dummy, Ded_{it} , equal to one if the regional government i introduces a deduction in (at least) one tax during the year¹⁸. $Left_{it}$ is a dummy variable equal to one if the party in office in a specific region and year is to the left of the political spectrum. Pop_{it} is the total population and accounts for regional size. At the same time, when regions acquired some tax power to modify those taxes ceded by the central government (in 1997 and 2002), a formal forum of interaction between each regional tax administration and the central one (bilateral nature), but also one of multilateral nature, was set up. This might have had an impact on the cooperation among regional tax administrations. However, it is not possible to identify its effect due to the absence of qualitative or quantitative information on the activity of these forums. Nevertheless, we implicitly account for this effect by including a set of time dummies, τ_t . We finally introduce fixed effects (ϑ_{ij}) to account for unobserved heterogeneity among CAs¹⁹, while ε_{ijt} is an idiosyncratic error that varies across time and pair of regions.²⁰ The parameters of Eq. (1) are estimated by maximum likelihood.

¹⁸ In our framework – in contrast with the hypothesis proposed by Bachetta and Espinosa (1995) – it is unlikely that a CA behaves strategically and lowers the tax burden via tax rate cuts so as to induce, to a certain measure, taxpayers to err in their tax returns: taxpayers would pay less and the CA would collect more tax revenues. All the same, in our case, it is difficult to identify such behaviour, since the information on the misreported tax revenues transmitted is available at an aggregated level and not tax by tax.

¹⁹ The “quality of the tax administration” would be an interesting control variable, as rightly suggested by a reviewer. However, this cannot be identified from our data, so we cannot identify its impact. The fixed effects should control for it as long as we suppose administrative quality does not vary much over time.

²⁰ In particular, $\vartheta_{ij} \sim N(0, \sigma_\vartheta)$ and $\varepsilon_{ijt} \sim N(0, \sigma_\varepsilon)$.

In order to have a better understanding of the determinants of the tax information sharing process, we extend this model in a dynamic fashion allowing for sluggish adjustment in the endogenous variable. It might take time for the regional tax authorities to process all the misreported tax revenues, so inertia might play a role in this process. Thus, following Wooldridge (2002, pp. 542-543), we also estimate a dynamic Tobit model with unobserved effects:

$$Trans_Rev_{ijt} = \max[0, \gamma g(Trans_Rev_{ijt-1}) + \delta Rec_Rev_{ijt-1} + Y_{ijt}\boldsymbol{\varphi} + X_{it}\boldsymbol{\rho} + \tau_t + c_{ij} + \epsilon_{ijt}]. \quad (2)$$

As in Eq. (1), we expect reciprocity to positively impact the cooperative behaviour of the regional tax authorities and then expect δ to be positive. In addition, we test the persistency hypothesis. In this regard, the function $g(\cdot)$ allows $Trans_Rev_{ijt-1}$ to appear in a variety of ways. We employ two alternative specifications:

- (i) $g(Trans_Rev_{ijt-1}) = Trans_Rev_{ijt-1}$; and
- (ii) $g(Trans_Rev_{ijt-1}) = \{1[Trans_Rev_{ijt-1} = 0]; 1[Trans_Rev_{ijt-1} > 0] \times Trans_Rev_{ijt-1}\}$, where $1[\cdot]$ is the indicator function.

The first approach is the standard dynamic model and, in this case, we expect γ to be positive; that is, cooperative behaviour in the previous period is expected to foster present cooperation. The second approach allows the effect of the lagged endogenous variable to be different depending on whether the previous response was a corner solution (zero) or strictly positive; then, in this case, γ is a vector 2×1 (see Wooldridge, 2002, pp. 542-543). Specifically, in this case, we expect to find a persistent behaviour over time, so that zero-valued transmitted misreported revenue in $t-1$ is expected to negatively impact the cooperative behaviour while the component $1[Trans_Rev_{ijt-1} > 0] \times Trans_Rev_{ijt-1}$ is expected to be positively related to the propensity to cooperate at time t .

In dynamic Tobit models with unobserved effects, the treatment of the initial observations is a key issue.²¹ Wooldridge (2005) proposes a fairly general and tractable solution to this econometric issue. This approach consists in specifying a distribution for the unobserved effect, c_{ij} , given the initial value, TR_{i0} , and the exogenous variables in all time periods. This leads to a fairly straightforward procedure that is no different from the standard static random-effects Tobit model. For practical purposes, the only difference between the exogenous initial values assumption and Wooldridge's approach is that the latter includes the initial values of the endogenous variable as additional explanatory variables in the regression.²²

²¹ The ideal case would be that the observed panel dataset starts together with the stochastic process. In this case, the initial values are known constants. If data is not collected at the beginning of the process, assuming that the initial values are exogenous, this might lead to bias and inconsistency in the estimators (Heckman, 1981; Hyslop, 1999; Honore, 2002). The first period in our dataset is 1989 but the decentralisation of the relevant taxes began in the mid-1980s, thus there are a few years for which this data is missing. Although the assumption of exogenous initial values might not be too strong because the missing years are relatively few in comparison to the extent of the dataset, the most appropriate approach is to assume that the initial values are endogenous. For a formal discussion of this issue see, for example, Akay (2009).

²² For a formal discussion of these issues and a formal derivation of this model, see Wooldridge (2002, pp. 542-543; 2005).

In our framework, the main incentives for a CA not to cooperate are the administrative costs and the financial costs of losing the financial yield of undue tax revenues. Thus, we suspect that a CA with relatively short-term budget constraints will decide to reduce cooperation. In order to identify the role of financial/budget constraints in influencing reciprocity, we interact Rec_Rev_{ijt-1} with $1[Def_{it}]$, a dummy equal to one if region i expects a deficit in period t . We perform this interaction for both the static and the dynamic models. Eq. (2) is then modified as follows:

$$Trans_Rev_{ijt} = \max[0, \gamma g(Trans_Rev_{ijt-1}) + \delta_1 Rec_Rev_{ijt-1} + \delta_2 Rec_Rev_{ijt-1} \times 1[Def_{it}] + Y_{ijt}\boldsymbol{\varphi} + X'_{it}\boldsymbol{\rho} + \tau_t + c_{ij} + \epsilon_{ijt}]. \quad (3)$$

Eq. (1) is also modified in a similar fashion. We expect δ_2 to be negative.

To conclude our empirical analysis, we investigate two additional and potentially important dimensions of heterogeneity in the effect of reciprocity on cooperation. First, we consider differences in the size across different regions by interacting Rec_Rev_{ijt-1} with $1[Pop_{it}]$, a dummy equal to one if region i has a population higher than the average. Thus Eq. (2) is modified in this way:

$$Trans_Rev_{ijt} = \max[0, \gamma g(Trans_Rev_{ijt-1}) + \pi_1 Rec_Rev_{ijt-1} + \pi_2 Rec_Rev_{ijt-1} \times 1[Pop_{it}] + Y_{ijt}\boldsymbol{\varphi} + X'_{it}\boldsymbol{\rho} + \tau_t + c_{ij} + \epsilon_{ijt}]. \quad (4)$$

Again, Eq. (1) changes in a similar way. We expect the reciprocity linkage to be weaker for bigger regions, since they have less to gain from reciprocity and thus we expect π_2 to be negative. Indeed, the size of the aggregate tax base, which is proxied by population, influences regional behaviour in a similar way to the one presented in the asymmetric competition literature (see, for example, Bucovetsky, 1991; Wilson, 1991). Namely, since bigger regions have a larger aggregate tax base and corresponding revenues, they obtain a smaller marginal benefit from reciprocal cooperation. We will test for this source of heterogeneity.

The second source of heterogeneity we want to investigate relates to the electoral cycle. We think that, in electoral years, tax authorities might want to end the administrative period with few pending information-sharing processes given the potential change of the executive. Thus, we expect the reciprocity linkage to be stronger in electoral years. Then, we interact Rec_Rev_{ijt-1} with El_{it} and Eq. (2) is modified in this way:

$$Trans_Rev_{ijt} = \max[0, \gamma g(Trans_Rev_{ijt-1}) + \mu_1 Rec_Rev_{ijt-1} + \mu_2 Rec_Rev_{ijt-1} \times El_{it} + Y_{ijt}\boldsymbol{\varphi} + X'_{it}\boldsymbol{\rho} + \tau_t + c_{ij} + \epsilon_{ijt}]. \quad (5)$$

Eq. (1) changes in a similar way. We expect μ_2 to be positive.

Data and Sources

The data on the cases of misreported taxes and their corresponding revenues, in addition to the regional tax and audit revenues and the dummy Ded_{it} , are extracted from the report entitled “Informe sobre la cesión de tributos a las Comunidades Autónomas”. The other variables are obtained from the following statistical sources. The distance between two CAs is the Euclidean distance between their capitals, is calculated using their geographical coordinates and is expressed in kilometres. The political alignment is defined using the information on the political colour of the governments in office, which we also employ for the definition of the variable $Left_{it}$. This information is obtained from Zarate’s Political Collections website (<http://zarate.eu/spain2.htm>). The relative GDP is based on data from the Spanish National Institute of Statistics (INE). The transfers-expenditure ratio is constructed as the ratio between the total amount of transfers received from the central government (extracted from the INE database) and the total regional expenditure (extracted from the Ministry of Economy and Finance database). The deficit is that expected at the beginning of the fiscal year, and is extracted from the database of the Ministry of Economy and Finance. The information on election years is obtained from the Ministry of the Interior’s website (<http://goo.gl/YCS3J>). In Table 1, we report the summary statistics.

Table 1. Summary Statistics

Variable	Measurement unit	Obs.	Mean	Std. Dev.	Min	Max
Transmitted Tax Revenues	thousands of 2001 euro	4,203	144.87	1,179.61	0	37,111.18
Received Tax Revenues	thousands of 2001 euro	4,206	114.30	954.11	0	38,900.90
Cases of Transmitted Misreported Taxes	number of cases	4,410	22.53	196.28	0	10,533
Cases of Received Misreported Taxes	number of cases	4,410	36.13	505.42	0	22,944
Distance	kilometres	4,410	630.73	512.75	31	2204
Political Alignment	dummy	4,410	0.51	0.50	0	1
Relative GDP	Ratio	4,410	1.04	0.29	0.46	2.15
Tot_Reg_Tax_Revenues	millions of 2001 euros	4,410	72.51	104.64	1.73	775.02
Tot_Reg_Audit_Revenues	millions of 2001 euros	3,990	3.59	6.69	0	49.85
Deficit	thousands of 2001 euro	4,200	-68,860.48	27,1390.3	-24,78177	1,270,978
1[Deficit]	dummy	4,200	0.38	0.49	0	1
Transfers/Expenditure	share of expenditure financed by transfers	4,410	0.35	0.17	-0.04	1.37
Leftist Government	dummy	4,410	0.44	0.50	0	1
Election Year	dummy	4,410	0.24	0.43	0	1
Deduction	dummy	4,410	0.15	0.35	0	1
Population	thousands of people	4,410	2,542.28	2,168.17	261.34	8,150.47

RESULTS

In Table 2, we present the results of the estimation of Eq. (1), that is, the static model. We report a GLS random-effects specification in column (1), a standard Tobit model in column (2), and column (3) reports the random-effects Tobit model, which is our preferred estimation strategy. The amount of misreported tax revenues transmitted by CA i to CA j positively depends on reciprocity, which is proxied by the time-lagged tax revenues received by CA i from CA j . This result is robust to the different specifications. According to the random effects Tobit model reported in column (3), a one euro increase in the tax revenues received by CA i from CA j in year $t-1$ results in an increase of 0.385 euros of tax revenues being transmitted from CA i to CA j in year t , holding all other variables constant.

Clearly, the amount of misreported revenues increases as the number of cases of transmitted misreported taxes grows. Specifically, according to model (3), one additional case of misreported taxes leads to an increase in transmitted revenues of almost 6.5 thousand euros, keeping constant all the other variables. The estimate of the distance between regions is significant and robust to the two different Tobit specifications presenting negative coefficients: two distant regions share less misreported revenues than is the case between two closer CAs. This corroborates previous results, as we saw in the literature review. Furthermore, we find that the deficit negatively impacts the cooperative behaviour of the tax administration. Those CAs with a higher expected deficit at the beginning of the year are less willing to transfer misreported tax revenues.

Table 2. Determinants of the Information-Sharing Process: TOBIT-RE and Alternative Specifications

Estimator	(1) GLS-RE	(2) TOBIT	(3) TOBIT-RE
L.Received Tax Revenues	0.467*** (10.456)	0.438*** (7.351)	0.385*** (6.311)
Cases of Transmitted Misreported Taxes	5.891*** (23.516)	6.892*** (20.554)	6.478*** (17.874)
Distance	-0.017 (-0.562)	-0.288*** (-5.850)	-0.299*** (-4.603)
Political Alignment	-64.845** (-2.094)	-61.081 (-1.293)	-45.212 (-0.880)
Relative GDP	-36.360 (-0.587)	-4.180 (-0.043)	14.190 (0.113)
Tot_Reg_Tax_Revenues	11.970 (0.717)	1.062 (0.042)	8.352 (0.295)
Tot_Reg_Audit_Revenues	-0.648 (-0.777)	-1.219 (-1.008)	-1.158 (-0.908)
Deficit	-0.000* (-1.848)	-0.000** (-2.052)	-0.000* (-1.768)
Transfers/Expenditure	161.385 (1.037)	396.833 (1.580)	366.767 (1.400)
Election Year	-2.153 (-0.061)	-73.051 (-1.340)	-74.340 (-1.212)
Deduction	-8.960 (-0.162)	9.324 (0.116)	0.885 (0.011)
Leftist Government	-12.665 (-0.180)	-113.368 (-1.126)	-89.040 (-0.846)
Population	0.006 (0.765)	0.065*** (5.545)	0.069*** (4.401)
_cons	48.804 (0.346)	-184.113 (-0.851)	-220.779 (-0.915)
<i>Observations</i>	3,446	3,446	3,446
<i>Censored Observations</i>	1,504	1,504	1,504
<i>Number of groups (couple of regions)</i>	210	210	210
R ²	0.244	-	-
Log likelihood	-	-17,134.759	-17,112.908
Wald chi2	1100.793	1036.608	785.558
p-value	0.0000	0.0000	0.0000

Notes: t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For all specifications, we report χ^2 statistics and p-values for the Wald test of joint significance. Time effects and regional dummies are included in all specifications.

As for the control variables, we find that regional size, proxied by population, is positively associated with the transfer of misreported tax revenues. None of the remaining covariates is found to be significant, but they are jointly statistically significant according to the Wald test.

In Table 3, we present the results of the estimation of the alternative specifications of Eq. (2) that we use to test the persistency hypothesis. In columns (1) and (2), we employ specification (i), while in columns (3) and (4) we use specification (ii).²³ The dynamic Tobit models in columns (2) and (4) are estimated by employing Wooldridge's (2005) approach, while the models in columns (1) and (3) are estimated by assuming exogenous initial values. The results suggest that there is a sluggish adjustment in the process of transmission of misreported tax revenues. In models (1) and (2), the coefficients of $Trans_Rev_{ijt-1}$ suggest that a one euro increase in misreported tax revenues transmitted by CA i to CA j in the previous year leads to an increase of almost 0.235 euros in the transmitted misreported revenues in the current year. Moreover, the results obtained by means of the estimation of models (3) and (4) corroborate our hypothesis of congruency in the behaviour of the regional tax authorities. The CAs that did not transmit revenues in $t-1$ tend to transmit less revenues in t , while the CAs that had transmitted revenues in $t-1$ transfer, on average, 0.023 euros more in t for any additional euro transmitted in $t-1$.

The initial value of the transmitted misreported revenues does not turn out to be significant, suggesting that there is no correlation between the unobserved heterogeneity and the initial condition. This is probably due to the fact that the first period in our panel dataset coincides mostly with the true starting point generating the process. Although Wooldridge's method is the most appropriate for the estimation of this process, this result indicates that the bias in the estimation of $g(Trans_Rev_{ijt-1})$ under the exogenous initial values assumption is not severe as confirmed by the magnitudes of the coefficients obtained through the two methodologies that are almost equal. Taking inertia into account, though, does not modify the main results obtained when estimating Eq. (1). In particular, reciprocity remains a driving force of the process.

²³ Specifically in columns (1) and (2) we set $g(Trans_Rev_{ijt-1}) = Trans_Rev_{ijt-1}$, while in columns (3) and (4) we assume $g(Trans_Rev_{ijt-1}) = \{1[Trans_Rev_{ijt-1} = 0]; 1[Trans_Rev_{ijt-1} > 0] \times Trans_Rev_{ijt-1}\}$.

Table 3. Determinants of the Information-Sharing process: Dynamic TOBIT-RE - Alternative Specifications

Estimator	(1) TOBIT-RE Exogenous initial values	(2) TOBIT-RE Wooldridge method	(3) TOBIT-RE Exogenous initial values	(4) TOBIT-RE Wooldridge method
L.Transmitted Tax Revenues	0.234*** (9.438)	0.235*** (9.456)	-	-
1[L.Transmitted Tax Revenues = 0]	-	-	-712.641*** (-13.257)	- (-13.168)
1[L.Transmitted Tax Revenues > 0] × L.Transmitted Tax Revenues	-	-	0.023*** (9.394)	0.023*** (9.393)
L.Received Tax Revenues	0.327*** (5.440)	0.327*** (5.442)	0.378*** (6.393)	0.378*** (6.393)
Transmitted Tax Revenues _{t=1989}	-	1.791 (1.512)	-	0.058 (0.063)
Cases of Misreported Taxes	5.926*** (16.620)	5.930*** (16.634)	5.848*** (17.365)	5.848*** (17.364)
Distance	-0.283*** (-4.718)	-0.271*** (-4.482)	-0.188*** (-3.745)	-0.187*** (-3.709)
Political Alignment	-33.301 (-0.664)	-32.097 (-0.640)	-66.883 (-1.403)	-66.807 (-1.401)
Relative GDP	10.889 (0.093)	17.485 (0.149)	-27.411 (-0.280)	-27.191 (-0.278)
Tot_Reg_Tax_Revenues	18.989 (0.665)	18.371 (0.644)	15.238 (0.576)	15.199 (0.574)
Tot_IGT_Audit_Revenues	-1.600 (-1.280)	-1.470 (-1.174)	-1.248 (-1.032)	-1.242 (-1.024)
Deficit	-0.000 (-1.485)	-0.000 (-1.459)	-0.000 (-1.502)	-0.000 (-1.500)
Transfers/Expenditure	446.910* (1.700)	434.541* (1.652)	376.867 (1.463)	376.357 (1.461)
Election Year	-49.016 (-0.824)	-51.271 (-0.862)	-40.768 (-0.739)	-40.883 (-0.741)
Deduction	-2.659 (-0.033)	-3.200 (-0.040)	-13.971 (-0.172)	-13.994 (-0.173)
Leftist Government	-96.980 (-0.944)	-98.314 (-0.957)	-77.896 (-0.777)	-77.954 (-0.778)
Population	0.065*** (4.468)	0.061*** (4.175)	0.038*** (3.193)	0.038*** (3.147)
_cons	-340.952 (-1.442)	-351.172 (-1.485)	-96.233 (-0.434)	-96.555 (-0.436)
Observations	3,405	3,405	3,405	3,405
Censored Observations	1,490	1,490.000	1,490	1,490
Number of groups (couple of regions)	210	210	210	210
Log likelihood	-16,845.972	-	-16,769.765	-16,769.763
		16,844.828		
Wald chi2	923.174	927.285	1,276.899	1,276.878
p-value	0.0000	0.0000	0.0000	0.0000

Notes: t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For all specifications, we report χ^2 statistics and p-values for the Wald test of joint significance. Time effects and regional dummies are included in all specifications.

In Table 4, we report the results of the estimation when we interact Rec_Rev_{ijt-1} with a dummy identifying periods of expected budget in deficit (Eq. 3).

Both in the static and in the dynamic approach, we still find reciprocity to be positively associated with the revenue transmission process, but this relationship is weaker during the periods in which CA i faces relatively more binding budget constraints. In the absence of deficit, the CAs transmit according to the different specifications at around 0.80 – 0.84 of every 1 euro received, while in the presence of (an expected) deficit, they transmit less than half that amount (0.29 – 0.35 of every 1 euro received).

Table 4. Determinants of the Information-Sharing Process: Interactions with High Deficit.

Estimator	(1) TOBIT-RE	(2) TOBIT-RE Wooldridge method	(3) TOBIT-RE Wooldridge method
L.Transmitted Tax Revenues	-	0.238*** (9.585)	-
1[L.Transmitted Tax Revenues = 0]	-	-	-704.264*** (-13.022)
1[L.Transmitted Tax Revenues > 0] × L.Transmitted Tax Revenues	-	-	0.023*** (9.482)
L.Received Tax Revenues	0.798*** (3.939)	0.816*** (4.113)	0.836*** (4.312)
L.Received Tax Revenues × 1[Deficit]	-0.442** (-2.125)	-0.525** (-2.570)	-0.495** (-2.474)
Transmitted Tax Revenues _{t=1989}	-	1.600 (1.372)	-0.090 (-0.098)
Cases of Misreported Taxes	6.492*** (17.969)	5.933*** (16.726)	5.829*** (17.329)
Distance	-0.296*** (-4.602)	-0.268*** (-4.497)	-0.185*** (-3.675)
Political Alignment	-47.302 (-0.923)	-34.418 (-0.689)	-67.460 (-1.417)
Relative GDP	25.821 (0.207)	30.349 (0.263)	-15.383 (-0.157)
Tot_Reg_Tax_Revenues	4.904 (0.173)	14.925 (0.525)	12.763 (0.482)
Tot_IGT_Audit_Revenues	-1.257 (-0.988)	-1.607 (-1.287)	-1.376 (-1.134)
1[Deficit]	74.688 (1.176)	77.504 (1.239)	54.397 (0.862)
Deficit	-0.000* (-1.764)	-0.000 (-1.448)	-0.000 (-1.294)
Transfer/Expenditure	301.820 (1.134)	371.018 (1.393)	333.402 (1.275)
Left	-73.329 (-1.197)	-50.600 (-0.854)	-41.794 (-0.756)
Election	3.812 (0.047)	0.536 (0.007)	-10.327 (-0.128)
Deduction	-70.078 (-0.662)	-78.229 (-0.758)	-62.124 (-0.614)
Population	0.069*** (4.400)	0.061*** (4.200)	0.038*** (3.158)
_cons	-256.972 (-1.065)	-392.334* (-1.660)	-133.715 (-0.600)
Linear Combinations			
L.Received Tax Revenues + L.Received Tax Revenues × 1[Deficit]	0.355*** (5.65)	0.291*** (4.69)	0.341*** (5.59)
<i>Observations</i>	3,446	3,405	3,405
<i>Censored Observations</i>	1,504	1,490	1,490
<i>Number of groups (couple of regions)</i>	210	210	210
Log likelihood	-17,110.207	-16,841.078	-16,766.540
Wald chi2	796.081	944.918	1,285.733
p-value	0.0000	0.0000	0.0000

Notes: *t* statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For all specifications, we report χ^2 statistics and *p*-values for the Wald test of joint significance. Time effects and regional dummies are included in all specifications.

Table 5 reports the results of the estimation of Eq. (4). By considering heterogeneity in regional size, we find that the reciprocity linkage is significantly weaker for more populated regions. In particular, smaller regions transmit 1.86 – 2.12 of every 1 euro received, while bigger regions transmit on average about one tenth of that amount (0.19 – 0.24 of every 1 euro received).

As a final exercise, we try to understand how the electoral cycle interacts with the reciprocity (Table 6). By interacting Rec_Rev_{ijt-1} with the electoral dummy, we find that reciprocity linkages are stronger in electoral years. In presence of elections, tax administrations transmit 2.41 – 2.52 of every 1 euro received while otherwise they transmit 0.27 – 0.33 of every 1 euro received.

Table 5. Determinants of the Information-Sharing Process: Interactions with High Population.

Estimator	(1) TOBIT-RE	(2) TOBIT-RE Wooldridge method	(3) TOBIT-RE Wooldridge method
L.Transmitted Tax Revenues	-	0.204*** (8.244)	-
1[L.Transmitted Tax Revenues = 0]	-	-	-682.449*** (-12.844)
1[L.Transmitted Tax Revenues > 0]×L.Transmitted Tax Revenues	-	-	0.020*** (8.163)
L.Received Tax Revenues	2.126*** (11.969)	1.871*** (10.536)	1.863*** (10.771)
L.Received Tax Revenues×1[HPop]	-1.902*** (-10.357)	-1.677*** (-9.174)	-1.620*** (-9.096)
Transmitted Tax Revenues _{t=1989}	-	1.326 (1.178)	-0.236 (-0.261)
Cases of Misreported Taxes	6.025*** (17.050)	5.600*** (16.050)	5.494*** (16.524)
Distance	-0.272*** (-4.401)	-0.252*** (-4.346)	-0.168*** (-3.364)
Political Alignment	-43.665 (-0.878)	-32.574 (-0.667)	-64.862 (-1.387)
Relative GDP	8.674 (0.073)	10.505 (0.094)	-23.210 (-0.241)
Tot_Reg_Tax_Revenues	23.738 (0.867)	32.390 (1.167)	27.615 (1.063)
Tot_IGT_Audit_Revenues	-0.747 (-0.604)	-1.124 (-0.921)	-0.938 (-0.789)
Deficit	-0.000 (-0.660)	-0.000 (-0.547)	-0.000 (-0.543)
Transfer/Expenditure	348.880 (1.372)	429.546* (1.674)	374.739 (1.484)
Left	-83.175 (-1.374)	-62.253 (-1.050)	-45.212 (-0.808)
Election	-26.256 (-0.335)	-26.948 (-0.344)	-37.095 (-0.467)
Deduction	-23.766 (-0.232)	-41.642 (-0.414)	-21.093 (-0.214)
Population	0.098*** (3.520)	0.088*** (3.331)	0.049** (2.177)
1[HPop]	-89.429 (-0.723)	-77.344 (-0.662)	9.864 (0.098)
_cons	-421.474* (-1.805)	-518.854** (-2.253)	-268.564 (-1.232)
Linear Combinations			
L.Received Tax Revenues +L.Received Tax Revenues× 1[HPop]	0.224*** (3.64)	0.194** (3.20)	0.244*** (4.08)
Observations	3,446	3,405	3,405
Censored Observations	1,504	1,490	1,490
Number of groups (couple of regions)	210	210	210
Log likelihood	-17060.233	-16803.221	-16729.211
Wald chi2	940.659	1060.800	1393.584
p-value	0.0000	0.0000	0.0000

Notes: *t* statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For all specifications, we report χ^2 statistics and *p*-values for the Wald test of joint significance. Time effects and regional dummies are included in all specifications.

Table 6: Determinants of the Information-Sharing Process: Interactions with Electoral Year.

Estimator	(1) TOBIT-RE	(2) TOBIT-RE Wooldridge method	(3) TOBIT-RE Wooldridge method
L.Transmitted Tax Revenues	-	0.242*** (9.937)	-
1[L.Transmitted Tax Revenues = 0]	-	-	-684.458*** (-12.918)
1[L.Transmitted Tax Revenues > 0]×L.Transmitted Tax Revenues	-	-	0.023*** (9.809)
L.Received Tax Revenues	0.333*** (5.595)	0.230*** (3.847)	0.274*** (4.654)
L.Received Tax Revenues×1[Election]	2.191*** (9.158)	2.196*** (9.353)	2.141*** (9.084)
Transmitted Tax Revenues _{t=1989}	-	1.553 (1.360)	-0.095 (-0.106)
Cases of Misreported Taxes	7.039*** (21.383)	6.114*** (17.574)	5.995*** (18.155)
Distance	-0.268*** (-5.553)	-0.250*** (-4.288)	-0.171*** (-3.457)
Political Alignment	-63.746 (-1.376)	-36.057 (-0.737)	-68.385 (-1.465)
Relative GDP	18.785 (0.198)	40.040 (0.354)	-2.340 (-0.024)
Tot_Reg_Tax_Revenues	-14.103 (-0.563)	-0.673 (-0.024)	-1.610 (-0.062)
Tot_IGT_Audit_Revenues	-1.519 (-1.279)	-1.729 (-1.411)	-1.544 (-1.297)
Deficit	-0.000* (-1.780)	-0.000 (-1.211)	-0.000 (-1.207)
Transfer/Expenditure	291.006 (1.179)	307.052 (1.192)	262.837 (1.039)
Left	-88.218* (-1.650)	-68.656 (-1.184)	-57.191 (-1.058)
Election	54.503 (0.687)	44.437 (0.568)	33.057 (0.416)
Deduction	-117.021 (-1.185)	-107.937 (-1.076)	-83.278 (-0.848)
Population	0.056*** (4.869)	0.052*** (3.665)	0.030** (2.556)
_cons	-117.984 (-0.555)	-261.440 (-1.134)	-29.716 (-0.137)
Linear Combinations			
L.Received Tax Revenues +L.Received Tax Revenues× 1[Election]	2.524*** (10.74)	2.426*** (10.47)	2.415*** (10.44)
Observations	3,446	3,405	3,405
Censored Observations	1,504	1,490	1,490
Number of groups (couple of regions)	210	210	210
Log likelihood	-17093.838	-16802.281	-16729.673
Wald chi2	1134.267	1061.395	1397.541
p-value	0.0000	0.0000	0.0000

Notes: *t* statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For all specifications, we report χ^2 statistics and p-values for the Wald test of joint significance. Time effects and regional dummies are included in all specifications.

CONCLUSIONS

We have analysed an area of horizontal tax interdependence that may occur in federal contexts, namely, the transmission of misreported tax revenues between sub-central tax administrations. We have obtained some evidence of the determinants of cooperation between the Spanish regional tax authorities. Our analysis suggests that cooperation is a matter of reciprocity and so we corroborate the results of the relevant theoretical literature. More specifically, the amount of tax revenues transmitted from one region to another positively depends on the revenues received from the latter in the previous period. This is the main result of the paper, and it is significant and robust to different specifications.

According to our results, the existence of a reciprocity linkage between CAs is crucial in determining their level of cooperation in managing misreporting of taxes. The robustness of this result suggests, even in absence of a counterfactual, that regional tax authorities behave strategically and do, under certain conditions, cooperate with each other in dealing with this problem. Namely, we have identified two main barriers that might reduce the sub-central tax administrations' incentives to cooperate. The existence of administrative and transaction costs directly related to the information-sharing process may induce a regional tax administration not to cooperate. Similarly, cooperation implies a financial cost due to the loss of revenue yields, which we find particularly important for the CAs that face budget constraints measured in terms of high deficit. However, our results suggest that. Typically, both barriers to cooperation act in the short-run, causing a delay in the transmission of relevant information. Indeed, the estimation of a dynamic model suggests that there is a sluggish adjustment in the setting of this process. We have explored different types of heterogeneity in the effect of reciprocity on the endogenous variable, finding that regional size reduces this effect on cooperation while the presence of electoral years tends to enhance it. Furthermore, we have found that the reciprocity link existing between two CAs becomes weaker when budget constraints are binding, *i.e.* in the presence of an expected deficit. This confirms the presence of barriers to cooperation due to the existence of financial costs.

Therefore, although a decentralised tax administration might entail some inefficiency costs due to a lack of incentives to collaborate across administrations, we prove that once they engage in cooperative behaviour, this is maintained, fostering even closer cooperation between them. This is a crucial point, because it suggests that once regional tax administrations become aware of the potential benefits of cooperation, they do not deviate from this equilibrium. In this regard, the central government could play a role in promoting the advantages of cooperation. All in all, this is good news for the functioning of a decentralised tax administration as, in this context, strategic considerations regarding tax base attraction might not be an issue. Further research, though, should be carried out to analyse other aspects of the functioning of tax administration in a federal context.

REFERENCES

- Akay, A. (2009). The Wooldridge Method for the Initial Values Problem Is Simple: What About Performance? IZA DP No. 3943.
- Bacchetta, P., & Espinosa, M. P. (1995). Information Sharing and Tax Competition Among Governments. *Journal of International Economics*, 39, 103–121.
- Bacchetta, P., & Espinosa, M. P. (2000). Exchange-of-Information Clauses in International Tax Treaties. *International Tax and Public Finance*, 7, 275–293.
- Bucovetsky, S. (1991). Asymmetric Tax Competition. *Journal of Urban Economics*, 30, 167-181.
- Chisik, R., & Davies, R. B. (2004). Asymmetric FDI and tax-treaty bargaining: theory and evidence. *Journal of Public Economics*, 88, 1119-1148.
- Cremer, H., & Gahvari, F. (2000). Tax evasion, fiscal competition and economic integration. *European Economic Review*, 44, 1633-1657.
- Dreher, A., Gehring, K., Kotsogiannis, C., & Marchesi S. (2014). Information transmission within federal fiscal architectures: Theory and evidence. CESifo Working Paper No. 4400.
- Durán-Cabré J. M., Esteller-Moré, A., & Salvadori, L. (2015). Empirical evidence on horizontal competition in tax enforcement. *International Tax and Public Finance*, 22(5), 834-860.
- Esteller-Moré, A. (2005). Is There a Connection Between the Tax Administration and The Political Power? *International Tax and Public Finance*, 12, 639-63.
- Esteller-Moré, A. (2008). Tax Administration in Spain: A Description and Analysis of the current institutional organization and some proposals of reform. In N. Bosch & J.M. Durán-Cabré (Eds.) *Fiscal Federalism and Political Decentralization: Lessons from Spain, Germany and Canada*, (pp. 209-247). Edward Elgar Publishing.
- Esteller-Moré, A. (2011). Is the Tax Administration Just a Money Machine? Empirical Evidence on Redistributive Politics. *Economics of Governance*, 12, 275-299.
- Elsayyad, M. (2012). Bargaining over tax information exchange. Max Planck Institute for Tax Law and Public Finance Working Paper No. 2012-02.
- Heckman, J.J. (1981). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process. In C. Manski and D. McFadden (Eds.) *Structural Analysis of Discrete Panel Data with Econometric Applications* (pp. 179-195). Cambridge, Mass: MIT Press.
- Hyslop, D.R. (1999). State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. *Econometrica*, 67, 1255-1294.
- Honore, B. (2002). Nonlinear Models with Panel Data. *Portuguese Economic Journal*, 1, 163–179.
- Huizinga H., & Nielsen, S.B., (2002). Withholding Taxes or Information Exchange: The Taxation of International Interest Flows. *Journal of Public Economics*, 87, 39–72.
- Johannesen, N., & Zucman, G. (2014). The End of Bank Secrecy? An Evaluation of the G20 Tax Haven Crackdown. *American Economic Journal: Economic Policy*, 6 (1): 65-91.

- Keen, M., & Ligthart, J.E. (2006a). Information sharing and international taxation: A primer. *International Tax and Public Finance*, 13, 81-110.
- Keen, M., & Ligthart, J.E. (2006b). Incentives and Information Exchange in International Taxation, *International Tax and Public Finance*, 13, 163-180.
- Ligthart J.E., & Voget, J. (2010). The determinants of Cross-Border Tax Information Sharing, mimeo.
- Martinez-Vazquez, J., & Timofeev, A. (2010) Choosing between centralized and decentralized models of tax administration. *International Journal of Public Administration*, 33, 601-619.
- Tanzi, V., & Zee, H.H. (2001). Can Information Exchange be Effective in Taxing Cross-Border Income Flows? In K. Andersson, P. Melz and C. Silfverberg (Eds.), *Modern Issues in the Law of International Taxation* (pp. 259-268). Stockholm: Kluwer Law International.
- Wilson, J. (1991). Tax Competition with Interregional Differences in Factor Endowments. *Regional Science and Urban Economics*, 21, 423-451.
- Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: MIT Press.
- Wooldridge, J.M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel-data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20, 39-54.

AN EXPLORATORY INVESTIGATION OF EXTRINSIC AND INTRINSIC MOTIVATIONS IN TAX AMNESTY DECISION-MAKING

Jonathan Farrar¹, Cass Hausserman²

Abstract

The tax compliance literature on tax amnesties does not explicitly consider the underlying motivational influences on taxpayers' self-correction decisions. Extant tax amnesty studies imply that extrinsic motives are the basis for self-correction, and only a few consider intrinsic motives (Rechberger, Hartner, Kirchler & Hämmerle, 2010; Torgler & Schaltegger, 2005). Consequently, we explore how extrinsic and intrinsic motives affect tax amnesty decision-making, following an unintentional taxpayer error. We conduct a quasi-experimental conjoint analysis on 1,266 taxpayers and vary the error magnitude. Results indicate that when taxpayers contemplate making a tax amnesty disclosure, desire to avoid a penalty is the most influential extrinsic motive, and responsibility to pay one's taxes is the most influential intrinsic motive. Extrinsic (intrinsic) influences account for about two-thirds (one-third) of the overall decision to make a tax amnesty disclosure. We also find that taxpayers' choices of extrinsic and intrinsic motives do not vary according to tax error magnitude. Implications for tax authorities and tax researchers are discussed.

INTRODUCTION

There is growing acceptance among tax researchers that taxpayer decision-making is complex and nuanced, and based on both extrinsic and intrinsic factors (Alm, Kirchler & Muehlbacher, 2012; Alm & Torgler, 2011; Feld & Frey, 2002). An individual who is motivated extrinsically expects to receive a benefit or avoid a punishment from an external source, whereas an individual who is motivated intrinsically is prompted to act for reasons of personal morality or internal feelings of satisfaction (Ryan & Deci, 2000). In the tax context, it follows that taxpayers may be motivated extrinsically or intrinsically to comply with a tax authority. From a tax authority perspective, appealing to taxpayers' intrinsic motivations rather than extrinsic motivations may be a lower-cost alternative, since it would not require the same human resources to detect and subsequently respond to a discovery of non-compliance. Given that tax authorities worldwide are increasingly facing budgetary restrictions, with 60% of tax authorities reporting reductions in staffing in recent years (OECD 2015a), developing a better understanding of taxpayers' intrinsic motivations relative to extrinsic motivations may be a useful objective.

Although tax researchers suggest that taxpayers have intrinsic motivations to comply with tax authorities (Alm, Kirchler & Muehlbacher, 2012; Alm & Torgler, 2011; Braithwaite, 2009; Dunn, Farrar & Hausserman, 2016; Frey, 1997), much remains to be understood about taxpayers'

¹ Associate Professor, Ted Rogers School of Management, Ryerson University

² Assistant Professor of Accounting, Portland State University

intrinsic motivations. As McKerchar, Bloomquist & Pope (2013, p.6) state, “Many have attempted to shed light on taxpayers’ internal motivations... but hard evidence is difficult to find.” Dwenger, Kleven, Rasul & Rincke (2016) acknowledge that relative to extrinsic motivations, intrinsic motivations are the hardest to measure and study empirically, and therefore the least well understood.

The purpose of this research is to develop a better understanding of extrinsic and intrinsic motivations in taxpayer decision-making. We develop this understanding by examining taxpayers’ attitudes towards making a tax amnesty disclosure, following an unintentional tax reporting error.³ Tax amnesties and voluntary disclosure programs are relatively low-cost compliance initiatives in which taxpayers are given the opportunity to self-correct errors on previously filed tax returns. A tax amnesty tends to be a one-time opportunity for self-correction with an expiry date, whereas voluntary disclosure programs are permanent and ongoing.⁴ By self-correcting, taxpayers pay the taxes that would have resulted had the amounts been correctly reported, usually with interest, but can avoid the penalties and/or sanctions that would have been imposed had the tax authority discovered the errors. Forty-seven countries now offer permanent amnesty programs (OECD 2015b), which suggests that tax authorities view the tax amnesty as an increasingly important tax compliance initiative.

Empirical studies suggest that amnesties have direct and indirect effects on tax revenues. Nevertheless, many studies suggest that tax amnesties are not particularly effective at encouraging participation in tax amnesties, resulting in net revenue gains from amnesty programs that are only modest at best. For example, Hasseldine (1998) analyzed a number of state tax amnesties in the United States, and found that amnesty revenues range from just 0.008 to 2 percent of state tax revenues. Moreover, these studies tend to assume that taxpayers consider only the economic costs and benefits of self-correction decisions (i.e. extrinsic factors), and therefore overlook non-economic factors (i.e. intrinsic factors) that may contribute to participation in tax amnesties. Perhaps tax amnesties would be more effective at generating revenue and increasing subsequent compliance if intrinsic motives were better understood and incorporated into the design of tax amnesties.

In this exploratory study, we address two primary research questions, as follows: 1) Which extrinsic and intrinsic motives have the greatest influence on taxpayers’ decisions to correct a tax error?; and 2) How does tax error magnitude impact these decisions? Since empirical research suggests that individuals justify dishonesty in small amounts, but less so in large amounts (Ariely 2008; Mazar, Amir & Ariely, 2008), we believe it is important to understand taxpayers’ motivations across different error thresholds. We use a quasi-experimental conjoint approach to investigate the relative importance of various extrinsic motives within a set of extrinsic motives, and the relative importance of various intrinsic motives within a set of intrinsic motives. Conjoint analysis is a statistical technique which determines how individuals choose among alternatives.

³ We focus on unintentional errors because the research is exploratory and we wanted to appeal to a broad base of taxpayers. Although tax authorities hope to encourage taxpayers who made unintentional *and* intentional errors to participate in amnesty programs, it would be unrealistic to ask participants questions about a scenario in which they were told to assume that they had purposefully evaded taxes.

⁴ In this article, we use the term ‘tax amnesty’ to refer to any program offered by a tax authority for taxpayers to self-correct past errors.

Participants were given combinations of extrinsic and intrinsic motives, and asked which combination would be most effective at convincing them to report their mistake. Our research is quasi-experimental, as we vary the error magnitude among participants in three ways (\$500, \$5,000, and \$50,000). Since the conjoint analysis methodology does not allow for examination of relative importance across factors, we also conduct a supplemental analysis to examine the relative importance of extrinsic versus intrinsic motivations. For this supplemental analysis, a different set of participants allocate points based on how important each of the eight factors would be in their decision, without regard to whether they are intrinsic or extrinsic.

We find that desire to avoid a penalty was the most important extrinsic motive, and responsibility for paying taxes owed was the most important intrinsic motive. Collectively, extrinsic (intrinsic) motives accounted for approximately 66% (34%) of the amnesty participation decision. We also find that taxpayers' relative preferences for extrinsic and intrinsic motives did not vary significantly according to error magnitudes. Thus, taxpayers' motivations for self-corrections appear stable, regardless of the size of their error.

This research contributes to the tax compliance literature by: identifying the most important extrinsic and intrinsic motives in a self-correction decision; by examining taxpayers' relative preferences within each set of motives; and by considering the role of error magnitude in compliance decisions. Existing research that investigates how intrinsic motives impact tax amnesties is limited (Rechberger et al., 2010; Torgler & Schaltegger, 2005), and empirical tax amnesty research tends to focus on subsequent income reporting or subsequent revenue collection effects, rather than on the reasons why taxpayers might be inclined to participate in a tax amnesty. In other words, tax amnesty research tends to be reactive rather than proactive, since it examines the after-effects of tax amnesties. In contrast, the present research contributes to this literature by examining taxpayer attitudes and intentions when an amnesty participation decision is contemplated. Our results should also be of interest to tax authorities looking to design or improve revenue collection through tax amnesties.

The remainder of the paper is organized as follows. In the next section, we conduct a literature review, followed by sections that describe our methodology and results, and discuss the implications of our findings for tax policy makers and tax researchers.

LITERATURE REVIEW

Universally, people are concerned with motivation, i.e., how to move themselves or others to act (Deci, 2016). While there are a number of possible theoretical frameworks that could be used to examine the tax amnesty decision, we explore this decision from the perspective of intrinsic and extrinsic motivations. A well-established psychology literature indicates that motivation can be either extrinsic (external to an individual, such as a third-party reward) or intrinsic (within an individual, such as self-esteem) (e.g., Deci, Koestner, & Ryan, 1999; Franco & Svensgaard, 2012; Ryan & Deci, 2000; Ryan, 2012; Sheldon & Kasser, 2008). Extrinsic and intrinsic motivations each affect economic decision-making (Gneezy, Meier, & Rey-Biel, 2011; Kakinaka & Kotani, 2011). Different regions of the brain are responsible for processing each type of motivation (Murayama, Matsumoto, Izuma, & Matsumoto, 2010).

The purpose of this literature review is to identify possible extrinsic and intrinsic motives that may be associated with tax amnesty decision-making. To this end, we peruse the empirical literatures on tax amnesties specifically, and tax compliance more broadly. We also searched the broader ethical decision-making and customer service literatures, since a decision to self-correct is an ethical decision, and involves a service interaction with a tax authority. Finally, we reviewed descriptions of past and current tax amnesty programs to isolate motives that could be relevant for tax amnesty decision-making.

Tax Amnesty Literature

Empirical studies have addressed two main ways in which amnesties can affect tax revenue collected: direct gains from participation in the amnesty (Alm and Beck, 1991; Fisher et al., 1989; Hasseldine, 1998; Luitel & Sobel, 2007), and indirect effects on tax compliance following an amnesty (Alm et al., 1990; Alm & Beck, 1993; Andreoni, 1991; Christian et al., 2002; López-Laborda & Rodrigo, 2003; Luna et al., 2006; Malik & Schwab, 1991; Rechberger et al., 2010; Torgler & Schaltegger, 2005; Young, 1994). This literature focuses on the impact of post-amnesty revenue collection, or taxpayers' compliance subsequent to an amnesty, rather than on taxpayers' underlying motivations to actually participate in an amnesty program. Nevertheless, the findings of each study are now briefly described, with a view to identifying possible motives that may be influential in the tax amnesty participation decision.

Fisher et al. (1989) examine the effectiveness of a state tax amnesty in Michigan, and find that overall revenues did not increase substantially due to the amnesty. Alm & Beck (1991) develop an economic model of amnesty participation, test it using data from twenty-eight states in the United States, and find that taxpayers disclose more in an amnesty program when probability of detection and penalties are expected to be greater. Hasseldine (1998) reviews 43 state tax amnesty programs in the United States, finds that tax amnesty revenues as a percentage of state tax revenues range from 0.008 to 2 percent, and finds that revenue collection declines with repeated amnesty programs. Luitel & Sobel (2007) find that states that offer repeated tax amnesties generate less revenue from the subsequent amnesties than the initial tax amnesties, and find reduced compliance following tax amnesties. Alm et al. (1990) also find that compliance decreases after an amnesty.

One economic model developed by Andreoni (1991) predicts that cheating increases when a permanent tax amnesty is enacted. Similarly, Malik & Schwab (1991)'s economic model shows that taxpayers report less income as the probability of an amnesty rises. Alm & Beck (1993) conduct a time-series economic analysis on a Colorado state tax amnesty, and suggest that this amnesty did not result in significant long-term post-amnesty revenue collection. Young (1994) examined characteristics of amnesty participants, and found that single males, and individuals with occupations in sales or who were self-employed, were more likely to participate in tax amnesties. Christian et al. (2002) found that the increase in tax revenues following a Michigan state tax amnesty was negligible. López-Laborda & Rodrigo (2003) evaluated the long-term impact of a Spanish tax amnesty, and found that the amnesty had no effect on tax collection in the short- or long-term. Torgler & Schaltegger (2005) experimentally examined the effect of taxpayers' voting approval for a tax amnesty on subsequent reporting compliance, and found that tax compliance increased only after group discussion and voting. Luna et al. (2006) review a number of state tax amnesties, identify features of each, and conclude that the long-term compliance effect is unclear.

Rechberger et al. (2010) examine the perceived justice of a tax amnesty on subsequent reporting compliance, and found that this relation is mediated by retribution and value restoration.

These studies tend to conclude that revenue collection effects of amnesties are modest at best, and that reporting compliance following an amnesty tends to suffer. Amnesties generate relatively little revenue in part because participation in amnesties is low. Our research focuses on antecedents of participation, which has the potential to improve revenue collection. Only one study (Alm & Beck, 1991) explicitly identifies possible motives that might influence taxpayer participation in tax amnesties. Alm & Beck (1991) conclude that probability of detection and penalties are important motives. For the majority of studies, the reason(s) taxpayers participate in tax amnesties is (are) not stated, and is (are) implied to be economic in nature. Consequently, much remains to be learned as to why taxpayers might be inclined to participate in a tax amnesty program.

Other Tax Compliance Literature

Consistent with the tax amnesty literature, conventional economic models of tax compliance suggest that the compliance decision is an economic decision, such that taxpayers weigh economic gains from evasion with possible sanctions from having their evasion detected and identified by the tax authority (e.g., Sandmo, 2005). In other words, the tax evasion decision is a function of detection likelihood, the size of the penalty, and the individual's degree of risk aversion (Slemrod, 2007). Thus, likelihood of detection, penalties, and risk tolerance are relevant for tax reporting compliance decisions, and may also be relevant for self-correction decisions.

A number of tax studies have also considered how tax morale impacts tax compliance (see Torgler, 2007, for a review). Tax morale is, "the collective name for all the non-rational factors and motivations – such as social norms, personal values, and various cognitive processes – that strongly affect an individual's voluntary compliance with laws" (Kornhauser, 2007, p. 602). Tax morale is synonymous with intrinsic motivation to pay taxes (McKerchar et al., 2013). The tax morale literature suggests that perceptions of fairness, trust in government, exchange equity, culture, and moral rules and sentiments all impact tax morale (e.g., Alm & Torgler, 2006; Frey & Torgler, 2007; McKerchar et al., 2013; Pope & McKerchar, 2011). Thus, these factors may also influence taxpayers' decisions to participate in a tax amnesty.

We also consulted tax compliance literature reviews to identify possible factors that may also be relevant for tax amnesty decision-making (Andreoni, Erard, & Feinstein, 1998; Cuccia, 1994; Fischer, Wartick, & Mark, 1992; Jackson & Milliron, 1986). Sanctions (penalties) and probability of detection were the predominant factors that were identified, along with other factors of guilt and social norms. All of these factors may affect taxpayers' extrinsic or intrinsic motivations. Other factors that affected tax compliance in these studies, such as demographic variables, are not inherently intrinsic or extrinsic, and thus we did not include them as possible extrinsic or intrinsic motives in the study, but rather measured and controlled for them when relevant.

Ethical Decision-Making and Customer Service Literatures

To identify other motives that may influence the tax amnesty decision, we examine the broader ethical decision-making and customer service literatures.⁵ We identified guilt, embarrassment and moral pride as motives relevant to ethical decision-making (Tangney, Steuwig & Mashek, 2007), as well as personal responsibility and peer reaction (Bobek, Hageman & Kelliher, 2013). Luria, Gal & Yagil (2009) identify belief that an individual will be treated fairly and ease of making restitution as additional factors that may influence individuals' willingness to report customer service complaints. Since taxpayers receive some degree of customer service when they interact with a tax authority, factors that influence customer service interactions may also be relevant for tax amnesty decisions.

Tax Amnesty Program Descriptions

We also read descriptions of tax amnesty programs worldwide (Baer & Le Borgne, 2008; Malherbe, 2011; OECD, 2015b) to identify other motives that could be relevant. Items that emerged related to: elapsed time; whether or not the taxpayer had the financial means to make restitution; the size of the penalty; and the amount of interest owing.

METHODOLOGY

Our research questions concern the relative importance of extrinsic and intrinsic motives to taxpayers when contemplating participation in a tax amnesty, and how their preferences for extrinsic and intrinsic motives vary according to the magnitude of their tax error. We address these questions using a quasi-experimental conjoint methodology. In the subsections that follow, we describe conjoint analysis, how we determined the extrinsic and intrinsic motives to use in the conjoint analysis, the experimental procedures, and the results.

Conjoint Analysis

In this section, we describe the conjoint analysis that we conducted to assess and understand the roles of extrinsic and intrinsic motives in tax amnesty decision-making. Conjoint analysis is a statistical technique, used most often in marketing research, to understand individuals' preferences for product features. A product has attributes (such as colour and size), and each attribute has several features (such as red, blue, and green colours; and small and large sizes). Conjoint analysis allows researchers to determine which combinations of product features are most preferred by consumers. Conjoint analysis helps researchers understand how consumers make choices among competing product features.

Conjoint statistical software computes a "part-worth utility" (a numerical value) for each feature of each attribute. Part-worth utilities of a particular product feature can be compared within each attribute to assess respondents' relative preferences of a product feature; and part-worth utilities from one attribute can be combined with part-worth utilities from another attribute, and compared with other similar combinations. Thus, in the above example, part-worth utilities from one colour

⁵ Using the ABI/Inform database, we searched ethics, hospitality, and marketing journals for the terms 'motivation' and 'motive'.

could be compared with part-worth utilities of all other colours; and a part-worth utility from a colour could be combined with a part-worth utility from a size, and compared to any other similar combination. However, part-worth utilities from one attribute cannot be compared with part-worth utilities from another attribute (Orme, 2010). Thus, in the above example, a part-worth utility of a colour could not be compared with a part-worth utility of a size.

While conjoint analysis tends to be used in marketing research, tax researchers have also employed this method, though not with respect to tax amnesty decisions (O'Neil, 1982; Blaufus, Bob, Hundsdoerfer, Kiesewetter & Weimann, 2013; Blaufus & Ortlieb, 2009; Hundsdoerfer & Sichtmann, 2009; Milliron & Toy, 1988; O'Neil, 1982). We examine the tax amnesty decision using a motivational psychology framework, in which there are two underlying attributes (extrinsic motivation and intrinsic motivation), with several features of each attribute (corresponding to specific extrinsic and intrinsic motives).

There are several advantages to using the conjoint methodology for this study. Conjoint analysis is a powerful way in which to analyze the relative importance of multiple features simultaneously. This method requires respondents to consider multiple attributes of their decision simultaneously, such that they must make trade-offs between different motives. Conjoint analysis also allows us to examine a larger number of motives than a traditional experiment, which is important, given the exploratory nature of the research. However, there is a restriction on the number of features each attribute can have; specifically, the number of features per attribute should not exceed four (Orme, 2010). Therefore, we are limited to including four extrinsic motives and four intrinsic motives in the conjoint analysis.

Selection of Extrinsic and Intrinsic Motives

To determine the four extrinsic and four intrinsic motives for use in the conjoint analysis, we initially compiled lists of 10 extrinsic and 11 intrinsic motives, using the findings from our literature review. Along with another academic, we independently categorized each motive we identified from the literature review as extrinsic or intrinsic, and reached a consensus as to the final classification.⁶

To validate our selection of motives, as well as our categorization of these motives as extrinsic and intrinsic, and to identify the four most influential motives across both categories, we conducted a pretest. Pretests are used commonly in behavioral research to substantiate an initial selection of items for use in a questionnaire, as well as to check for glitches in wording (e.g., Hite, 1998; Libby & Thorne, 2007).

The pretest was conducted on 65 adult students (average age of 28.8 years) in two tax classes.⁷ Participants read a brief background about tax amnesties, were given a list of motives, and were

⁶ We were unable to reach consensus as to the extrinsic or intrinsic nature of two factors (the amount of the mistake, and the amount of time that has passed since the mistake). To determine whether a typical taxpayer thought these would be intrinsic or extrinsic motivations, we surveyed participants in our pretest (see footnote 8). Regardless, this issue is not relevant, as neither factor was retained for the conjoint analysis.

⁷ The sample does not need to be the same as the population, as long as there is nothing in the sample that is expected to bias the results (Elliott, Hodge, Kennedy, & Pronk, 2007). Nothing about the pretest sample was expected to bias the results.

asked to rank them in terms of their importance in the tax amnesty decision.⁸ They were then asked if there were any additional reasons that were not listed that would be influential in this decision, and if any wording or any factors were unclear, which helped us to further refine the wording used, and helped to ensure that we had not overlooked other important motives. No additional motives were identified.

The pretest results (not tabulated) indicated that the most important extrinsic motives were: *wanting to avoid a penalty; the size of the penalty; probability of the error being detected by the tax authority; and effort to disclose.*⁹ There was also agreement that the most important intrinsic motives were: *responsibility to pay the taxes owed; satisfaction for correcting the mistake; feeling guilty for not paying the tax; and concern about how the taxpayer would be treated.*

Below, we tabulate the initial lists of motives, classified as extrinsic or intrinsic, and indicate in bold font the four extrinsic and four intrinsic motives retained for our subsequent conjoint study.

⁸ To cross-validate the pretest results, one class was given a list of all possible motives, while the other class was given lists of extrinsic and intrinsic motives separately. Using two groups of students provides greater assurance regarding the consistency of their rankings of the most important motives. We compared results from both classes, which were largely consistent (see footnote 9), i.e., the most important extrinsic and intrinsic motives were ranked similarly across both classes. At the end of the pretest, participants were also provided with a definition of intrinsic and extrinsic motivation, and asked whether they thought two ambiguous items (*the amount of the mistake* and *the amount of time that has passed since the mistake*, were primarily intrinsic or primarily extrinsic.

⁹ One class rated *amount of the mistake* just higher than *effort to disclose*, but we retained *effort to disclose*, since the size of the penalty is a function of the amount of the mistake, and the two items could be confounded.

Table 1 – Initial list of extrinsic and intrinsic motives when contemplating a tax amnesty disclosure *

	Extrinsic motives	Intrinsic motives
1)	Concern about other people's opinion of me for not paying the tax.	Feeling guilty about not paying the tax.
2)	The amount of time that has passed since the mistake.	Feeling guilty about making the mistake.
3)	The length of time that the voluntary disclosure program is available.	Feeling embarrassed for not paying the tax.
4)	The size of the penalty.	Feeling embarrassed for making the mistake.
5)	Wanting to avoid paying a penalty.	Feeling embarrassed if the mistake is discovered by the tax authority.
6)	The amount of interest.	Feeling satisfaction for paying the tax.
7)	The amount of the mistake.	Feeling satisfaction for correcting the mistake.
8)	Concern that paying the amount owed will affect my lifestyle.	Feeling responsibility to pay the taxes owed.
9)	The amount of effort required to disclose my mistake.	Feeling responsibility to pay taxes in general.
10)	The chance that I'll get caught in the future if I don't admit my mistake now.	Feeling uncertainty about being detected by the tax authority.
11)		Concern about how the tax authority will treat me if I admit my mistake.

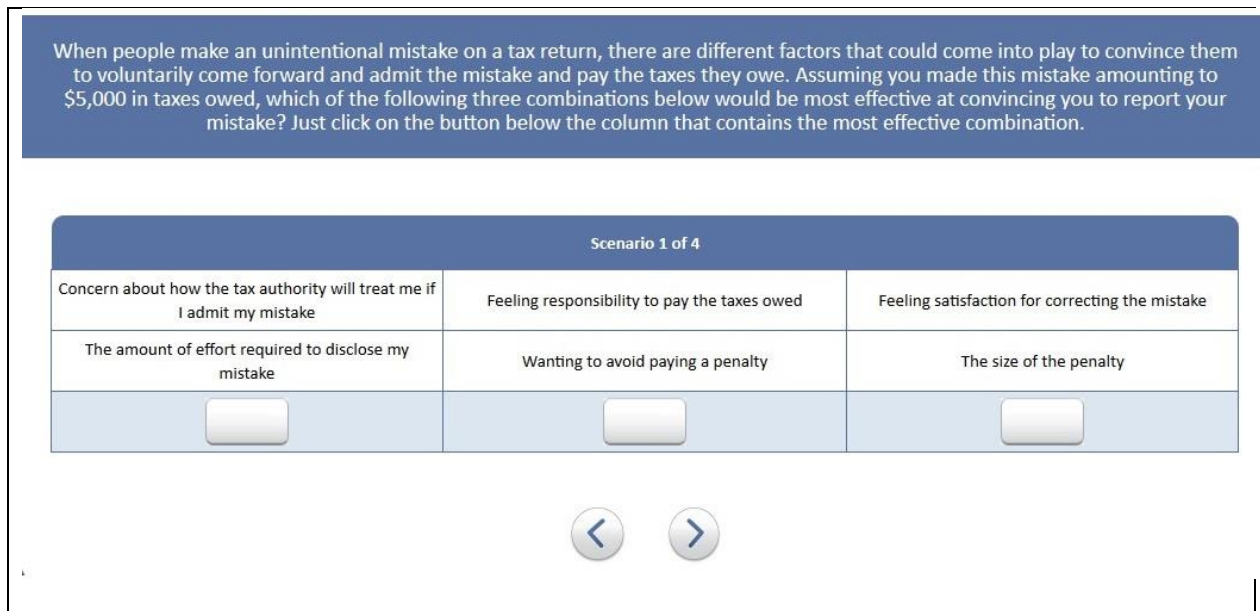
** Note: items in bold font represent items retained for use in the conjoint analysis.*

Procedures

We perform a choice-based conjoint analysis, using the shortlists of four extrinsic and four intrinsic motives, to gauge the relative importance of each motive within their respective motivation categories. Our design is a fractional-factorial design, in which selections of combinations of motives are presented to the respondents. A full-factorial design, in which all combinations are presented to participants, is impractical due to respondent fatigue, so fractional-factorial designs are used instead, and are just as effective as full-factorial designs (Tovares, Boatwright, & Cagan, 2014).

Respondents were United States taxpayers recruited from a large market research firm, and chosen randomly from across the United States, but segmented according to gender and age (individuals at least 18 years of age). Respondents read a brief background about tax amnesty programs, followed by a vignette in which they were asked to imagine that they had made an unintentional mistake on their tax return (either \$500, \$5,000, or \$50,000), and were then presented with a series of screens that presented three choice combinations per screen. Each choice combination had one extrinsic and one intrinsic motive.¹⁰ Respondents were asked which combination of the three would be most effective at convincing them to report their mistake to the tax authority. Their answer determined, in part, which combinations appeared on the next screen, as programmed by the software. A sample screenshot is below in Figure 1.

Figure 1 – Sample screenshot



¹⁰ We acknowledge that tax amnesty decisions are not comprised of exactly one intrinsic and one extrinsic motive, but in order to determine the relative importance of the intrinsic and extrinsic motives using conjoint analysis, this set-up is required. To address the concern that the decision may not be based on one extrinsic and one intrinsic motive, we conducted a supplemental analysis, in which participants freely indicated the importance of each motive without regard to whether it was extrinsic or intrinsic.

It is important to examine how taxpayers' motivational preferences may vary across different error thresholds. Ariely (2008) and Mazar, Amir, & Ariely (2008) suggest and find evidence that individuals can justify dishonesty in small amounts without compromising a positive view of themselves, but not in large amounts. Consequently, taxpayers with relatively small errors may be motivated differently than taxpayers with relatively large errors.

We were unable to find any guidance in the academic literature on choices of dollar magnitudes for use in an experiment. Although our choices of dollar magnitudes are subjective, we used the vignette development suggestions of Weber (1992) and Hughes & Huby (2004), who emphasize that vignettes must be as realistic as possible. We chose an upper limit of \$50,000, after consulting an industry publication which reports dollar amounts of frauds, as well as considering anecdotal evidence reported in the business press, and used our pretest to verify that this amount was plausible to respondents.¹¹ Once we chose the upper limit, we chose the other two amounts (\$5,000 and \$500) as equidistant intervals on a logarithmic scale.

A total of 1,266 taxpayers completed the instrument. To ensure high data quality, the instrument contained two 'attention check' questions.¹² Respondents who failed one or both of these questions were terminated, and their responses were not included in the final tally. The average age of a respondent was 45.3 years, and 52.1% of the sample was female. Detailed demographic information is contained in Table 2.

To gauge the effectiveness of the error magnitude manipulation, respondents were asked to rate their agreement with the following statement: *The amount of taxes owed was quite large*. Respondents rated this statement on a 7-point Likert scale, where 1=strongly disagree and 7=strongly agree. The mean scores for respondents in the \$500, \$5,000, and \$50,000 conditions were 3.87, 5.80, and 6.61, respectively, which are in the expected direction. Furthermore, Mann-Whitney U-tests showed significant differences in these scores between the \$500 and \$5,000 groups ($Z=15.53$, $p<0.01$), and between the \$5,000 and \$50,000 groups ($Z=10.61$, $p<0.01$). Therefore, the error magnitudes were effectively manipulated across conditions.

¹¹ We consulted the "Report to the Nation" of the Association of Certified Fraud Examiners (ACFE 2014), which reported a median dollar amount for intentional mistakes of \$145,000. Our upper threshold of \$50,000 is well below this median figure, as we felt that taxpayers would have difficulty relating to any higher amount as an unintentional mistake. Furthermore, our upper threshold of \$50,000 appears plausible, given stories in the American popular press of two potential government appointees who made unintentional tax errors in the amounts of \$34,000 and \$140,000 (Reuters, 2009). None of the pretest participants expressed concern over an upper limit of \$50,000.

¹² One question was, "In the scenario, how much did it say you owe in taxes?" Respondents could choose between a) \$500, b) \$5,000, or c) \$50,000. The correct answer depended upon experimental condition. The other question was, "In the scenario, what was the reason provided for why you owe taxes?" The options were: a) you intentionally made a mistake in the past; b) you unintentionally made a mistake in the past; and c) the IRS made a mistake, and as a result, you owe more in taxes.

Table 2 – Demographic profile statistics

	\$500 error	\$5,000 error	\$50,000 error	All responses
Sample size	423	422	421	1,266
Gender				
<i>male</i>	197 (46.6%)	211 (50.0%)	199 (47.3%)	607 (47.9%)
<i>female</i>	226 (53.4%)	211 (50.0%)	222 (52.7%)	659 (52.1%)
Age	43.9	44.8	47.4	45.3
Education				
<i>less than high school</i>	3 (0.7%)	3 (0.7%)	7 (1.7%)	13 (1.0%)
<i>high school</i>	75 (17.7%)	71 (16.8%)	71 (16.9%)	217 (17.1%)
<i>some college courses</i>	129 (30.5%)	140 (33.2%)	154 (36.6%)	423 (33.4%)
<i>college graduate</i>	151 (35.7%)	140 (33.2%)	122 (29.0%)	413 (32.6%)
<i>post-graduate degree</i>	65 (15.4%)	68 (16.1%)	67 (15.9%)	200 (15.8%)
Income				
<i>less than \$25,000</i>	75 (17.7%)	80 (19.0%)	95 (22.6%)	250 (19.7%)
<i>\$25,000 to \$49,999</i>	126 (29.8%)	107 (25.4%)	107 (25.4%)	340 (26.9%)
<i>\$50,000 - \$74,999</i>	85 (20.1%)	89 (21.1%)	82 (19.5%)	256 (20.2%)
<i>\$75,000 - \$99,999</i>	72 (17.0%)	59 (14.0%)	53 (12.6%)	184 (14.5%)
<i>Over \$100,000</i>	57 (13.5%)	80 (19.0%)	79 (18.8%)	216 (17.1%)
<i>Prefer not to respond</i>	8 (1.9%)	7 (1.7%)	5 (1.2%)	20 (1.6%)
Ever made a tax amnesty disclosure?				
<i>yes</i>	16 (3.8%)	26 (6.2%)	11 (2.6%)	53 (4.2%)
<i>no</i>	407 (96.2%)	396 (93.8%)	410 (97.4%)	1,213 (95.8%)

RESULTS

We first examined which extrinsic and intrinsic motives have the greatest influence on taxpayers' decisions to correct a tax error (Research Question 1). To address this question, we examined the part-worth utilities of each motive. We then take the antilog of the part-worth utilities in order to express them as a proportion, so that we can predict the percentage of the population that is influenced by each attribute (Sawtooth Software, 2002). The average from all respondents of the part-worth utilities, the antilogs, and the relative percentages are presented in Table 3.

Table 3 – Part-worth utility scores for extrinsic and intrinsic motives

Factor	\$500 error			\$5,000 error			\$50,000 error			MANOVA between groups	
	Part-worth utility	<i>antilog</i>	Percentage	Part-worth utility	<i>antilog</i>	Percentage	Part-worth utility	<i>antilog</i>	Percentage		
EXTRINSIC											
Avoiding a penalty	0.92	2.50	51.76%	0.94	2.57	52.76%	0.98	2.68	53.67%	F=1.04, p=0.35	
Future detection	0.04	1.04	21.53%	0.02	1.02	21.02%	0.03	1.03	20.65%	F=0.04, p=0.96	
Penalty size	-0.18	0.83	17.18%	-0.21	0.81	16.58%	-0.15	0.86	17.20%	F=1.26, p=0.28	
Effort	-0.78	0.46	9.53%	-0.76	0.47	9.64%	-0.86	0.42	8.48%	F=1.75, p=0.17	
			100.0%				100.0%				100.0%
INTRINSIC											
Factor	\$500 error			\$5,000 error			\$50,000 error			MANOVA between groups	
	Part-worth utility	<i>antilog</i>	Percentage	Part-worth utility	<i>antilog</i>	Percentage	Part-worth utility	<i>antilog</i>	Percentage		
Responsibility	0.54	1.72	39.36%	0.56	1.76	39.45%	0.61	1.84	41.25%	F=0.90, p=0.41	
Satisfaction	0.26	1.30	29.75%	0.32	1.38	31.05%	0.27	1.32	29.42%	F=0.60, p=0.55	
Guilt	-0.27	0.76	17.47%	-0.23	0.80	17.92%	-0.23	0.79	17.68%	F=0.61, p=0.54	
Treatment	-0.53	0.59	13.42%	-0.66	0.52	11.58%	-0.65	0.52	11.66%	F=1.02, p=0.36	
			100.0%				100.0%				100.0%

Results indicate that the most important extrinsic factor is the desire to avoid paying a penalty, followed by the probability of future detection, the size of the penalty, and the amount of effort required to disclose the mistake. Overall, the desire to avoid paying a penalty was rated to be approximately 2.5 times more important than the next most important factor, the probability of future detection, and was rated just over three times more important than the size of the penalty.

Results also indicate that the most important intrinsic factor is a feeling of responsibility to pay the taxes owed, followed by satisfaction for correcting the mistake, feeling guilty for not paying the tax, and concern for how the authorities would treat the taxpayer. The responsibility factor was rated only slightly higher than the satisfaction factor (about 0.15 times), but responsibility was rated more than twice as important as guilt and almost three times as important as treatment by the tax authority.

We then examined how tax error magnitudes impacted respondents' choices of extrinsic and intrinsic motives (Research Question 2). To address this question, we conducted a MANOVA of the part-worth utilities for all 8 motives across each error condition. As the columns in Table 3 show, there were no significant differences in part-worth utilities for any extrinsic or intrinsic factors. Therefore, the relative importance of any extrinsic or intrinsic motives did not vary significantly by error condition.

Supplemental Analysis

Because conjoint analysis does not allow us to examine the relative importance of intrinsic versus extrinsic motives, we conducted a supplemental analysis to address this issue. Using a different set of participants (also recruited from an online survey company), we presented 299 participants with the same background information and scenario as in the conjoint analysis study.¹³ Rather than asking them to select among pairs of intrinsic and extrinsic motives, we asked them to allocate 100 points to each of the eight possible motives, based on how influential the motives would be if making a tax amnesty decision. As in the conjoint study, we split the participants into three groups, according to three error magnitudes (\$500, \$5,000, and \$50,000). Overall, we found that participants allocated 66% of their points to extrinsic factors and 34% of their points to intrinsic factors. These findings suggest that extrinsic factors are significantly more important than intrinsic factors.¹⁴

Similar to the conjoint study, using MANOVA, we did not find any significant differences in extrinsic/intrinsic allocations across error conditions at the 0.05 level of significance.¹⁵ Therefore, this finding provides additional independent evidence that taxpayers' motivations appear stable across error magnitudes. Table 4 reports the mean extrinsic and intrinsic scores for this supplemental analysis, across the three error conditions.

¹³ The average age of a respondent was 36.9 and 54% were male.

¹⁴ Wilcoxon signed-rank tests across all three conditions were significant: in the \$500 condition, $Z=-5.841$, $p<0.01$; in the \$5,000 condition, $Z=-6.407$, $p<0.01$; and in the \$50,000 condition, $Z=-5.383$, $p<0.01$.

¹⁵ The amount of points allocated to any motive did not vary significantly across any error condition at the 0.05 level of significance. Furthermore, Mann-Whitney U test results are as follows: for the scores in the \$500 vs. \$5,000 condition, $Z=-0.623$, $p=0.53$; for the scores in the \$5,000 vs. \$50,000 condition, $Z=-0.819$, $p=0.41$; and for the scores in the \$500 vs. \$50,000 condition, $p=0.87$.

Table 4 – Allocation between extrinsic and intrinsic motives

	All data (n=299)	\$500 error (n=102)	\$5,000 error (n=100)	\$50,000 error (n=97)
Intrinsic only	34.16	34.46	32.34	35.71
Extrinsic only	65.84	65.54	67.66	64.29
	100	100	100	100

DISCUSSION & IMPLICATIONS

In this research, we provide preliminary evidence that both extrinsic and intrinsic motives influence taxpayer decision-making in a tax amnesty context. Specifically, we identify and analyze the key extrinsic and intrinsic motives that influence taxpayers' amnesty decisions. We first developed shortlists of the four most important extrinsic and intrinsic motives, respectively. We then used both sets of motives in a quasi-experimental conjoint analysis, where we were able to determine respondents' relative preferences for each motive within both categories of motivation, across three error magnitudes (\$500, \$5,000, and \$50,000). We found that desire to avoid a penalty was the most important extrinsic motive, and responsibility for paying the taxes owed was the most important intrinsic motive. Our results further indicate that the magnitude of the tax error does not influence the relative importance of extrinsic or intrinsic motivational factors in tax amnesty decision-making. Thus, taxpayers' motivational preferences appear stable across tax error magnitudes. In a supplemental analysis, we determined that extrinsic factors are responsible for approximately two-thirds of the decision to participate in a tax amnesty, whereas intrinsic factors are responsible for approximately one-third.

We extend and contribute to the literature on tax amnesties by identifying influential extrinsic and intrinsic motives, and by showing that intrinsic motivations have an influential albeit less significant role than extrinsic motivations on taxpayer's decisions to participate in a tax amnesty. We also extend the broader tax compliance literature by finding that in a tax amnesty context, the size of taxpayers' errors does not seem to influence their underlying motivations to cooperate with a tax authority.

As with all behavioral research, this study has limitations. To prevent decision fatigue, and because of the constraints of conjoint analysis, the list of potential motives was not exhaustive. Therefore, it is possible that our results would have differed had we used additional motives. Also, since this study was tested on taxpayers from the United States, results should be applied cautiously to other jurisdictions. Future research could consider how taxpayers in other countries are motivated to make amnesty disclosures, since there may be cultural differences that impact taxpayers' extrinsic and intrinsic motivations. We also acknowledge that participants in our study were asked about their motivations in a hypothetical scenario rather than in an actual situation. Thus, our study captures participants' intended, rather than actual, motivations.

Although most of the respondents do not have direct experience with tax amnesties, it would be infeasible to recruit respondents who had participated, or would consider participating, in a tax amnesty. However, participants in this study were able to relate to and understand the scenario, evidenced by correctly answering the attention check questions, and from reading the comments in the pretest. Additionally, hypothetical vignettes are a useful tool when studying

an ethical topic with which individuals may not be personally familiar (Dunn et al., 2016; Weber, 1992; Hughes & Huby, 2004; Mudrack & Mason, 2013; Weber, 1992).

Tax compliance researchers have suggested that a responsive regulation approach between taxpayers and tax authorities (Braithwaite & Braithwaite, 2001), which seeks to foster cooperative attitudes from taxpayers using intrinsic motivations in addition to extrinsic motivations, may be more effective than a traditional deterrence approach, which relies on extrinsic motivations. Furthermore, Kirchler (2007), Kirchler, Hoelzl, & Wahl (2008), and Alm et al. (2012) suggest a “slippery slope framework” of tax compliance, in which voluntary compliance and enforced compliance are both present. According to this framework, voluntary compliance depends on the right mix of trust in tax authorities, which is largely a function of intrinsic motivations, such as perceived fairness, and enforcement, such as threat of penalties. The results from our conjoint analysis provide support for this paradigm and additional insight regarding specific factors that are relevant in a tax amnesty decision.

Both our conjoint study and supplemental analysis, using different samples, revealed that taxpayers’ intrinsic and extrinsic motivations did not significantly differ across error magnitudes. This finding may suggest that it is a taxpayer's anxiety about having made an error that triggers specific motivational responses, rather than the amount of the error. This suggestion is consistent with Bobek, Hatfield & Wentzel (2007), who found that taxpayers perceive satisfaction and enjoyment from receiving refunds, to the extent that they will overpay interim tax payments to ensure they are in a refund position when they submit their annual tax return. In Bobek et al. (2007), it was the fact that taxpayers were in a refund position, rather than the size of the refund, that provided satisfaction. Similarly, the reasons for correcting an error may not be related to the magnitude of an error, but rather to the fact that a taxpayer has anxiety over making an error.

Another implication of our findings, specific to tax amnesties, is that authorities may be most likely to encourage taxpayers’ participation in an amnesty if taxpayers know that they can avoid penalties, since this was the most important extrinsic motive in the conjoint study. Thus, tax authorities could promote tax amnesties with a message that focuses on penalty avoidance. Since tax authorities worldwide are increasingly adopting permanent amnesty programs (OECD, 2015b), promoting awareness of these initiatives will become increasingly important, as will tailoring a message to encourage taxpayers to self-correct.

Another implication of our findings is that intrinsic motivations appear to have some role in taxpayers’ amnesty decision-making. Although the role of intrinsic motivations does not appear to be as influential as extrinsic motivations, it may be possible for tax authorities to appeal to intrinsic motives, which may simultaneously enhance taxpayers’ extrinsic motivations and, in turn, increase the likelihood of cooperation with authorities. Specifically, Frey & Jegen (2001) suggest that, in some circumstances, intrinsic motivations can enhance the strength of extrinsic motivations in influencing behaviour (Frey & Jegen, 2001). While further research is needed to examine the joint influence of extrinsic and intrinsic motives on the decision to participate in a tax amnesty, given the relative lack of success of tax amnesty programs which rely solely on economic (extrinsic) motives, tax authorities may want to consider how influential intrinsic motives, such as responsibility to pay one’s taxes, can be paired with influential extrinsic motives, such as desire to avoid penalties, to enhance the effectiveness of a tax amnesty program. We encourage further research to investigate this possibility.

REFERENCES

- Association of Certified Fraud Examiners (ACFE). 2014. *Report to the Nations on Occupational Fraud and Abuse*. Austin, TX: ACFE.
- Alm, J., McKee, M., & Beck, W. (1990). Amazing grace: Tax amnesties and compliance. *National Tax Journal*, 43(1), 23-37.
- Alm, J., & Beck, W. (1991). Wiping the slate clean: Individual response to state tax amnesties. *Southern Economic Journal*, 57(4), 1043-1053.
- Alm, J., & Beck, W. (1993). Tax amnesties and compliance in the long run: A time series analysis. *National Tax Journal*, 46(1), 53-60.
- Alm, J., & Torgler, B. (2006). Tax morale and culture differences in the United States and Europe. *Journal of Economic Psychology*, 27(3), 224-246.
- Alm, J., & Torgler, B. (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics*, 101(4), 635-651.
- Alm, J., Kirchler, E., & Muehlbacher, S. (2012). Combining psychology and economics in the analysis of compliance: From enforcement to cooperation. *Economic Analysis & Policy*, 42(2), 133-151.
- Andreoni, J. (1991). The desirability of a permanent tax amnesty. *Journal of Public Economics*, 45(1), 143-159.
- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *The Journal of Economic Literature*, 36(2), 818-860.
- Ariely, D. (2008). *Predictably Irrational: The Hidden Forces that Shape our Decisions*. New York: Harper Collins.
- Baer, K., & Le Borgne, E. (2008). *Tax Amnesties: Theory, Trends, and Some Alternatives*. Washington, DC: International Monetary Fund.
- Blaufus, K., & Ortlieb, R. (2009). Is simple better? A conjoint analysis of the effects of tax complexity on employee preferences concerning company pension plans. *Schmalenbach Business Review*, 61(1), 60-83.
- Blaufus, K., Bob, J., Hundsdoerfer, J., Kiesewetter, D., & Weimann, J. (2013). Decision heuristics and tax perception - An analysis of a tax-cut-cum-base-broadening policy. *Journal of Economic Psychology*, 35(1), 1-16.
- Bobek, D., Hatfield, C., & Wentzel, K. (2007). An investigation of why taxpayers prefer refunds: A theory of planned behavior approach. *Journal of the American Taxation Association*, 29(1), 93-111.
- Bobek, D., Hageman, A., & Kelliher, C. (2013). Analyzing the role of social norms in tax compliance behavior. *Journal of Business Ethics*, 115(3), 451-468.
- Braithwaite, V., & Braithwaite, J. (2001). An evolving compliance model for tax enforcement. In Shover, N. & Wright, J. P. (eds.), *Crimes of Privilege: Readings in White-Collar Crime* (pp. 1-19). New York: Oxford University Press.
- Braithwaite, V. (2009). *Defiance in taxation and governance*. Cheltenham, UK: Edward Elgar Publishing.
- Christian, C., Gupta, S., & Young, J. (2002). Evidence on subsequent filing from the state of Michigan's income tax amnesty. *National Tax Journal*, 55(4), 703-721.

- Cuccia, A. (1994). The economics of tax compliance: What do we know and where do we go? *Journal of Accounting Literature* 13, 81-116.
- Deci, E., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668.
- Deci, E. (2016). Overview of Self-Determination Theory. Online: <http://selfdeterminationtheory.org/theory/>
- Dunn, P., Farrar, J., & Hausserman, C. (2016). The influence of guilt cognitions on taxpayers' voluntary disclosures. *Journal of Business Ethics*, forthcoming.
- Dwenger, N., Kleven, H., Rasul, I., & Rincke, J. (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy*, forthcoming.
- Elliott, W.B., Hodge, F., Kennedy, J., & Pronk, M. (2007). Are MBA students a good proxy for nonprofessional investors? *The Accounting Review*, 82(1), 139-168.
- Feld, B., & Feld, L. (2002). Deterrence and morale in taxation: An empirical analysis. CESifo Working paper no. 760.
- Fischer, C., Wartick, M., & Mark, M. (1992). Detection probability and taxpayer compliance: A review of the literature. *Journal of Accounting Literature*, 11, 1-46.
- Fisher, R., Goddeeris, J., & Young, J. (1989). Participation in tax amnesties: The individual income tax. *National Tax Journal*, 42(1), 15-27.
- Franco, J., & Svensgaard, A. (2012). *Handbook on psychology of motivation: New research*. Hauppauge, NY: Nova Science Publishers.
- Frey, B. (1997). *Not just for the money: An economic theory of personal motivation*. Cheltenham, UK: Edward Elgar Publishing.
- Frey, B.S., and Jegen, R. (2001). Motivation-crowding theory. *Journal of Economic Surveys*, 15(5): 589-611.
- Frey, B., & Torgler, B. (2007). Tax morale and condition cooperation. *Journal of Comparative Economics*, 35(2), 136-159.
- Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25(4), 191-210.
- Hasseldine, J. (1998). Tax amnesties: An international review. *Bulletin for International Fiscal Documentation*, 52(7), 303-310.
- Hite, P. (1998). An examination of factors influencing financial reporting decisions of small business owner-managers. *Behavioral Research in Accounting*, 10(2), 159-178.
- Hughes, R., and Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review*, 11(1), 36-51.
- Hundsdoerfer, J., & Sichtmann, C. (2009). The importance of taxes in entrepreneurial decisions: an analysis of practicing physicians' behavior. *Review of Managerial Science*, 3(1), 19-40.
- Jackson, B., & Milliron, V. (1986). Tax compliance research: Findings, problems, and prospects. *Journal of Accounting Literature* 5, 125-161.

- Kakinaka, M., & Kotani, K. (2011). An interplay between intrinsic and extrinsic motivations on voluntary contributions to a public good in a large economy. *Public Choice*, 147(1-2), 29-41.
- Kirchler, E. (2007). *Economic psychology of tax behaviour*. New York: Cambridge University Press.
- Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The “slippery slope” framework. *Journal of Economic Psychology*, 29(2), 210-255.
- Kornhauser, M. (2007). A tax morale approach to compliance: Recommendations for the IRS. *Florida Tax Review*, 8(6), 599-640.
- Libby, T. & Thorne, L. (2007). The development of a measure of auditors’ virtue. *Journal of Business Ethics*, 71(1), 89-99.
- López-Laborda, J., & Rodrigo, F. (2003). Tax amnesties and income tax compliance: The case of Spain. *Fiscal Studies*, 24(1), 73-96.
- Luitel, H., & Sobel, R. (2007). The revenue impact of repeated tax amnesties. *Public Budgeting & Finance*, 27(3), 19-38.
- Luna, L., Brown, M., Mantzke, K., Tower, R., & Wright, L. (2006). State tax amnesties: Forgiveness is divine – and possibly profitable. *State Tax Notes*, August 21, 497-511.
- Luria, G., Gal, I., & Yagil, D. (2009). Employees’ willingness to report service complaints. *Journal of Service Research*, 12(2), 156-174.
- Malherbe, J. (2011). *Tax Amnesties*. Amsterdam, The Netherlands: Kluwer Law International BV.
- Malik, A., & Schwab, R. (1991). The economics of tax amnesties. *Journal of Public Economics*, 46(1), 29-49.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(December), 633-644.
- McKerchar, M., Bloomquist, K., and Pope, J. (2013). Indicators of tax morale: An exploratory study. *eJournal of Tax Research*, 11(1), 5-22.
- Milliron, V., & Toy, D. (1988). Tax compliance: An investigation of key features. *The Journal of the American Tax Association*, 9(1), 84-104.
- Mudrack, P., and E. S. Mason (2013). Dilemmas, conspiracies, and Sophie’s choice: Vignette themes and ethical judgments. *Journal of Business Ethics* 118(5), 639-653.
- Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America*, 107(49), 20911-20916.
- OECD 2015a. Tax administration 2015: Comparative information on OECD and other advanced and emerging economies. Paris: OECD Publishing. Online: http://www.keepeek.com/Digital-Asset-Management/oecd/taxation/tax-administration-2015_tax_admin-2015-en#page1
- OECD 2015b. Update on voluntary disclosure programmes: A pathway to tax compliance. OECD, August 2015. Retrieved from <http://www.oecd.org/ctp/exchange-of-tax-information/Voluntary-Disclosure-Programmes-2015.pdf>

- O'Neil, C. (1982). The targeted jobs credit: An evaluation of its impact on the employment decision process. *Journal of the American Taxation Association*, 3(2), 15-22.
- Orme, B. (2010). *Getting started with conjoint analysis: Strategies for product design and pricing research* (2nd ed.). Madison, WI: Research Publishers.
- Pope, J., & McKerchar, M. (2011). Understanding tax morale and its effect on individual taxpayer compliance. *British Tax Review*, 5, 587-601.
- Rechberger, S., Hartner, M., Kirchler, E., & Hämmerle, F. (2010). Tax amnesties, justice perceptions, and filing behavior: A simulation study. *Law & Policy*, 32(2), 214-225.
- Reuters (2009). Health nominee Sibelius reveals tax errors. Online: <http://www.reuters.com/article/us-healthcare-usa-sebelius-tax-idUSTRE53004X20090401>
- Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.
- Ryan, R. (2012). *Oxford handbook of human motivation*. Oxford: Oxford University Press.
- Sandmo, A. (2005). The theory of tax evasion: A retrospective view. *National Tax Journal*, 58(4), 643-663.
- Sawtooth Software (2002). Conjoint Value Analysis (CVA) Version 3.0 Technical Paper Series. Retrieved from <https://www.sawtoothsoftware.com/download/techpap/cva3tech.pdf>
- Sheldon, K., & Kasser, T. (2008). Psychological threat and extrinsic goal striving. *Motivation and Emotion*, 32(1), 37-45.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *The Journal of Economic Perspectives*, 21(1), 25-48.
- Tangney, J., Steuwig, J., & Mashek, D. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58, 345-372.
- Torgler, B., & Schaltegger, C. (2005). Tax amnesties and political participation. *Public Finance Review*, 33(3), 403-431.
- Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis*. Cheltenham, UK: Edward Elgar Publishing.
- Tovares, N., Boatwright, P., & Cagan, J. (2014). Experiential conjoint analysis: An experience-based method for eliciting, capturing, and modeling consumer preference. *Journal of Mechanical Design*, 136(October), 1-12.
- Weber, J. (1992). Scenarios in business ethics research: Review, critical assessment, and recommendations. *Business Ethics Quarterly*, 2(2), 137-160.
- Young, J. (1994). Factors associated with noncompliance: Evidence from the Michigan Tax Amnesty program. *Journal of the American Taxation Association*, 16(2), 82-105.

A GLOBAL FRAMEWORK ON THE FORMULATION AND IMPLEMENTATION OF INTERNATIONAL TAXATION STANDARDS

*Satoru Araki*¹

Abstract

With the emergence of the Group of Twenty (G20) as the premier international forum, international taxation standards have attracted interest from the perspective of international financial law or governance. This article aims to analyse the global process of formulating and implementing international taxation rules and standards, and attempts to propose a new global framework model which includes a range of ‘players’, from international organisations to national governments, who are involved in the international taxation system. It has been found that the global architecture or framework of international taxation standards is still in the process of evolution. The primary focus is to make the global framework inclusive of emerging and developing countries. On the other hand, as the capacity of the OECD and its frameworks as a technical assistance provider is not necessarily sufficient, a second focus for the global framework is to become part of functional partnerships which enable international organisations to address issues faced by tax authorities, particularly those in developing countries, effectively. Lastly, in order for the framework to have a truly global reach, there will need to be a solid network of regional frameworks or forums.

1. INTRODUCTION

Cross-border taxation issues are regulated using international tax rules and standards; whilst the classic example is a tax treaty that resolves double-taxation issues, there are also rules and standards relating to matters such as transfer pricing, dispute resolution mechanisms and tax information exchange.²

To what extent are these international taxation rules and standards different to domestic taxation rules? Tax policy is a sovereign issue, and domestic tax policy is in the hands of each national government. On the other hand, international taxation, including tax treaties, regulates and adjusts cross-border taxation issues, and policy co-ordination between countries is often, if not always, required. Discussions surrounding the international tax system often date back to the first bilateral tax agreement between Austria-Hungary and Prussia in 1899 (Jogarajan, 2011).

In order to smooth international taxation issues, it is desirable to have international models or practices which many countries desire to follow. The earliest of these international models dates back to bilateral conventions drafted as part of the work of the League of Nations in the 1920s (United Nations, 2001). In as far as tax policy is a sovereign issue, international standards or practices are not directly enforceable, nor are they automatically incorporated into national legal systems; on the other hand, these standards or practices set by international frameworks have a considerable influence on policy and administration in national governments, and in terms of global influence, they have a wider reach than laws enacted by a

¹ Former Public Management Specialist, Asian Development Bank in Metro Manila.

² In English case law, which had held that a tax claim made by a foreign state was not enforceable in England, *Government of India v Taylor* [1955] AC 491 (House of Lords), required tax administration bodies to work together for the international exchange of tax information.

single government. These models, guidelines or practices are often called ‘soft law’. Although it is difficult to define soft law as an alternative concept to hard law: it is not directly binding, and is elaborated with the aim of imposing standards of conduct that are not enforceable through traditional means of force (Sarmiento, 2010). Brummer (2015) also argues that, in contrast to coercive hard law, the essence of soft law is an expression of co-operation and the production of dominant norms for the co-ordination of behaviour.

As international rules and standards set through international organisations have global influence, it is worth examining the process of formulating and implementing these rules and standards, which is, in a sense, no less complicated than that of domestic legislation. Indeed, the implementation and monitoring of international rules and standards through domestic legislation in each government constitute part of a broader global process and framework for international rules and standards. Furthermore, as will be discussed in Section 2, the last several years have seen significant changes in the global process and framework due to the rise of emerging countries and the Group of Twenty (G20) as a body to endorse global agenda, and a comprehensive review of international taxation rules under the Organisation for Economic Co-operation and Development (OECD) Base Erosion and Profit Shifting (BEPS) project.

2. THE CURRENT LANDSCAPE

Currently, the OECD, an international organisation based in Paris, is a driving force for shaping rules and standards regarding international taxation. The OECD’s *Model Tax Convention on Income and on Capital* serves as a basis for more than 3500 bilateral tax treaties, along with *United Nations Model Double Taxation Convention between Developed and Developing Countries* (OECD, 2015a). Another signature product of the OECD in the international taxation field is its *Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations*, which are referred to as international standards of transfer pricing taxation. Furthermore, the Global Forum on Transparency and Exchange of Information for Tax Purposes (Global Forum) is a framework where OECD and non-OECD member jurisdictions are working together to promote standards in the international exchange of tax information between tax administration bodies.

What is the OECD’s decision-making procedure? Taxation falls within the remit of one of its specialised committees, the Committee on Fiscal Affairs (CFA), which consists of senior tax officials representing OECD member countries. As taxation is such a broad topic, the CFA delegates technical discussion to its subsidiary working parties, each of which comprises tax officials specialising in a particular issue, such as Working Party No. 1, which is in charge of tax treaties (Ault, 2009).

The global financial crisis of 2007-2008, combined with the rise of emerging economies represented by the new acronym of BRICS (Brazil, Russia, India, China and South Africa), has changed the landscape of the global political economy, which led to the designation of the G20, rather than the G7, as ‘the premier forum for international economic co-operation’ (G20 Leaders Statement: The Pittsburgh Summit, 2009). In a similar vein, business process in the OECD’s tax work metamorphosed from 2009, after the emergence of the G20. In the pre-G20 era, discussions were basically held between OECD member countries, predominantly advanced European countries; on the other hand, in the G20 era, emerging economies are invited to the process of making international standards on an equal footing with OECD member countries. The OECD’s CFA attaches importance to broader participation from emerging and developing countries in the implementation of these standards.

At the G20 London Summit in April 2009, in the context of tax information exchange, G20 leaders were ‘committed to make it easier for developing countries to secure the benefits of a new cooperative tax environment’ (Declaration on Strengthening the Financial System – London Summit, 2009). In September 2009, the OECD’s then framework for dialogues between OECD member and non-member jurisdictions was restructured as the current Global Forum, which opened its membership to all jurisdictions committed to implementing international standards on tax information exchange (OECD, 2009b).

As another symbol of the G20 era, the OECD’s BEPS project commenced and has proceeded under the auspices of the G20. The *G20 Leaders Declaration* of the Los Cabos Summit in June 2012, which stated, ‘We [G20 leaders] reiterate the need to prevent base erosion and profit shifting and we will follow with attention the ongoing work of the OECD in this area,’ was probably the first communication to put the buzzword BEPS into the public arena. When the OECD released the final output package of the BEPS project in October 2015, it was emphasised that, not only OECD member countries, but a wide range of countries, as well as other stakeholders, had directly participated in the decision-making process at the OECD’s CFA or had been consulted via various channels and events. In addition to the 34 official member countries of the OECD, other emerging and developing G20 member countries, including Argentina, Brazil, India, Indonesia, the People’s Republic of China, Russia, Saudi Arabia and South Africa, have directly participated in discussions at the OECD’s CFA with respect to the BEPS project (OECD, 2015a, p.4). Furthermore, the OECD has organised regional network meetings for dialogues with a broader group of developing countries and 59 countries had participated in these regional network meetings by the middle of 2016.³ These developments contrast with the pre-G20 era as, apart from the transparency and information exchange project (which focused on offshore financial centres), interaction with developing countries was chiefly conducted through ‘outreach’ activities, which aimed to disseminate the OECD’s products, such as *Model Tax Convention on Income and on Capital*, typically through one-week seminars (Ault, 2009, p.761).

This shift from the pre-G20 era to the G20 era with respect to the process of forming international standards in international taxation issues is illustrated in the following figure:

Figure 1: Shift of the Process of Forming International Standards

	Pre-G20 era	⇒	G20 era
Underlying international forum	G7	⇒	G20
Decision-making process	OECD member countries	⇒	OECD member and emerging countries
Relationship with developing countries	Dissemination of products through ‘outreach’ activities	⇒	Invite inputs to standard-making process through regional network activities

³ OECD. Developing countries and BEPS (Regional Networks on the BEPS Project). Retrieved from <http://www.oecd.org/tax/developing-countries-and-beps.htm#regionalnetworks>

3. MODELS OF THE PROCESS OF FORMING INTERNATIONAL RULES AND STANDARDS

This section will look at existing literature on the process of forming international rules and standards. Slaughter (2004) proposes the concept of transgovernmental networks where government officials in particular fields, such as financial regulation, collaborate internationally in order to address common issues, and identifies three broad categories of activities undertaken by these transgovernmental networks: information sharing, enforcement co-operation and rule harmonisation.

Drezner (2007) categorises international governmental organisations by three forms of membership: universes, clubs and neighbourhoods. In the context of setting international codes and standards in financial regulation in the 1990s and 2000s (i.e. the pre-G20 era), Drezner argues that club international organisations, the members of which largely consist of developed countries, such as the OECD, Basel Committee on Banking Supervision and Financial Action Task Force (FATF),⁴ have played greater roles than universal-membership international financial institutions (IFIs), such as the International Monetary Fund (IMF) and World Bank Group. Reaching consensus in universal-membership IFIs, where developed and developing countries have diverse positions and interests, has proved more challenging (Drezner, 2007, p.122).

Haas (1992) proposes the concept of an epistemic community as a network of professionals with recognised expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within that domain. In the context of European Union law, Schrauwen (2010) points out that an international expert group comprised of national officials, which meet and exchange expertise on a regular basis, could lead to a knowledge community which has influence over policy direction at the European Union or national levels.

More recently, Brummer (2015) explains global financial architecture, which sets international rules and standards in financial regulation, and ensures their implementation at national level, largely through four entities: 1) Agenda Setters; 2) Standard Setters; 3) Implementers; and 4) Monitors.

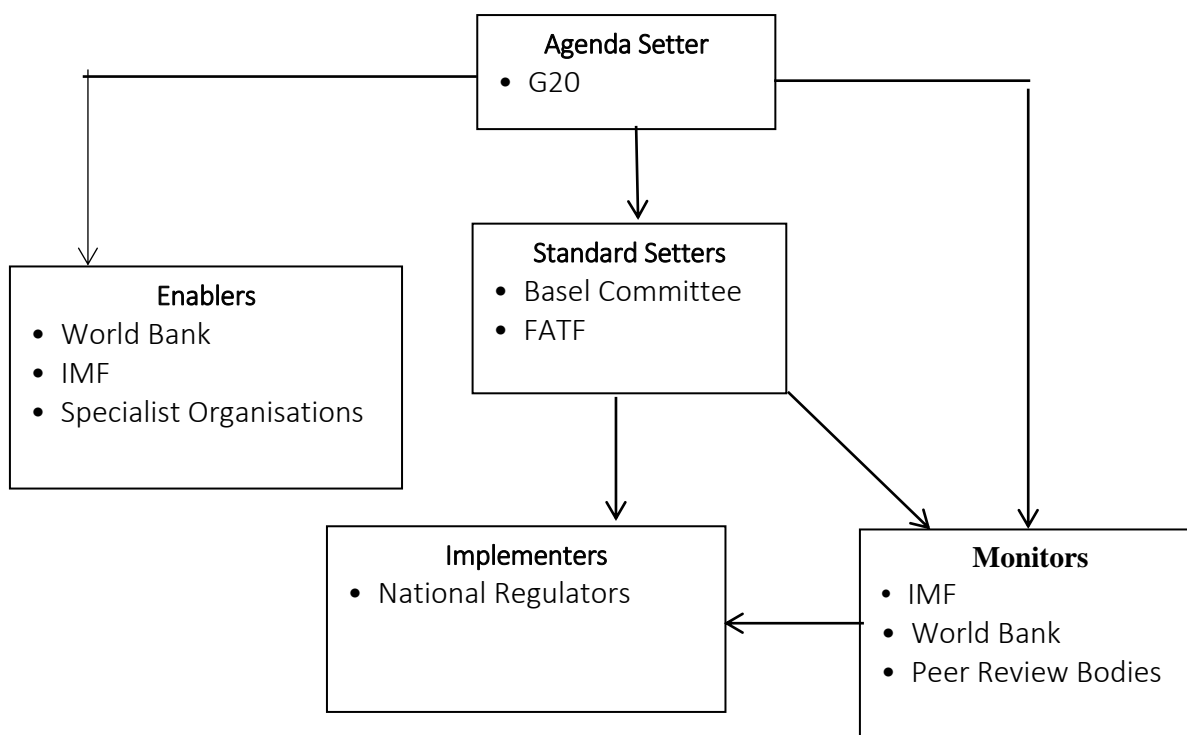
‘Agenda Setters’, which show policy objectives and directions, and provide ‘Standard Setters’ with political legitimacy, are represented by the G20, as well as by the Financial Stability Board. International standards called for by the ‘Agenda Setters’ are developed by ‘Standard Setters’ specialising in a particular field. The concept of the ‘Standard Setters’, examples of which include the OECD, Basel Committee on Banking Supervision and FATF, coincides with that of the club international organisations proposed by Drezner. Needless to say, international standards are expected to be implemented and enforced by national governments and public authorities, as ‘Implementers’. As a fourth function, the implementation of the international standards by national governments is to be monitored and ensured. This monitoring process is resource-intensive and requires the savvy of capacity development in developing countries. Whilst monitoring is often conducted as a peer review process by the members of a ‘Standard Setting’ organisation, Brummer names the IMF and World Bank as the most appropriate international bodies for the monitoring task; the two high-profile international financial

⁴ The objectives of the FATF, established in 1989, are to set and promote standards for combating money laundering and terrorist financing, and its secretariat is located within the OECD in Paris.

institutions (IFIs) have universal membership, as Drezner points out, and a larger capacity than 'Standard Setting' organisations, which have relatively small secretariats.

Grinberg (2015) proposes a modified version of Brummer's model of global financial architecture, in the context of international taxation, as shown in Figure 2.

Figure 2: Architecture of International Financial Law: Grinberg's Model



One of the main differences between Brummer's model and Grinberg's is that Grinberg adds a new category, 'Enablers' who, typically in developing countries, help national regulators (as 'Implementers') to meet international standards through technical assistance activities. The need for technical assistance can be identified through the monitoring process. In the context of international taxation, Grinberg cites the example of international standards in tax information exchange. The OECD's Global Forum promulgates the international standards, and conducts monitoring and peer review activities. On the other hand, the OECD, including the Global Forum, does not appear to have sufficient capacity and experience of working in developing countries (Asian Development Bank, 2014; Araki, 2015; Grinberg, 2015). Therefore, international organisations with more experience of development issues, such as the IMF and World Bank Group, can play a greater role in helping developing countries to implement international standards.

4. THE PLAYERS OF FRAMEWORK ON INTERNATIONAL TAXATION STANDARDS

Who, according to Brummer and Grinberg's models, are the players in the context of forming and implementing standards in international taxation? As discussed in Section 2 on the current

landscape, there is no question that the G20 is the premier international forum underlying the direction of work on international taxation, particularly since 2009 (Grinberg, 2015).

On the other hand, there is some question as to whether the G20 literally sets the agenda as a 'top-down' decision-making process; for it is unlikely that a substantial discussions on taxation are regularly held amongst the heads of states. The reality could be that agendas on international taxation are still proposed, if not set, by the OECD's Committee on Fiscal Affairs (CFA), which consists of senior government officials in charge of taxation.

In the case of the OECD's project on transparency and exchange of information for tax purposes, the G20 London Summit in April 2009 thrust a spotlight onto the project. Whilst the OECD had been working on the transparency of so-called tax havens since 1996 (Araki, 2015), it was not until March 2009, during the run-up to the London Summit, that some financial centres, such as Andorra, Lichtenstein and Monaco, expressed their commitment to implementing the OECD's standard on transparency and exchange of information (Global Forum, 2009, p.17). The G20's participation in, or support of, the project undoubtedly accelerated the emergence of a global system for tax information exchange.

With respect to the BEPS project, as cited in Section 2, the *G20 Leaders Declaration* of the Los Cabos Summit in June 2012 stated, 'We reiterate the need to prevent base erosion and profit shifting and we will follow with attention the ongoing work of the OECD in this area' (emphasis added). The BEPS project did not start from scratch, and the OECD had worked for several years on the aggressive tax planning project which served as a forerunner of the BEPS project.⁵ For example, in February 2011, the OECD published a report titled *Tackling Aggressive Tax Planning through Improved Transparency and Disclosure*, and the idea of the report was succeeded by the BEPS project's Action 12 on requiring taxpayers to disclose their aggressive tax planning arrangements (OECD, 2015b). When the BEPS final package (consisting of final reports on 15 actions) was released in 2015, it was submitted by the Secretary-General of the OECD to G20 leaders for endorsement in Antalya in November 2015 (G20 Leaders' Communiqué, 2015). The function of the G20 with respect to international taxation standards could be described more as an 'agenda-endorser' or 'approver', much like a board of directors, rather than as an author of a project proposal.

On the other hand, whilst the premier international forum has shifted from the G7 to the G20, the G7 still continues to deliver messages on international taxation issues in a proactive manner. *The Leaders' Declaration* of the G7 Ise-Shima Summit in May 2016 had a section on tax and transparency, and 'request[ed] to the OECD to establish objective criteria to identify non-cooperative jurisdictions with respect to tax transparency', which reflected public attention to offshore financial centres provoked by the so-called Panama Papers (OECD, 2016b).

With respect to 'Standard Setters', as mentioned in Section 2, the OECD's CFA is leading the development of standards in international taxation, whilst the United Nations' Committee of Experts on International Cooperation in Tax Matters is also carrying out the international taxation agenda, chiefly from developing countries' perspectives.⁶ When it comes to implementation, there is no difference from other regulatory issues: international rules and

⁵ A senior advisor at the OECD Centre for Tax Policy and Administration who has been involved in the BEPS project recalls that, when the BEPS' work started in 2012, the OECD had worked on aggressive tax planning for a few years. Cf. Russo, R. (2016).

⁶ Cf. United Nations Economic and Social Council. (2004, November, 11). *ECOSOC Resolution 2004/69: Committee of Experts on International Cooperation in Tax Matters*.

standards are not effective until they are implemented by national governments.⁷ It should be noted that, in many countries, taxation is dealt with by the treasury or ministry of finance in charge of tax policy and an inland revenue body in charge of tax administration.⁸

Does the international taxation field have separate frameworks for monitoring? For example, in the cases of combating money laundering and the financing of terrorism (AML/CFT), the assessment of compliance with international standards promulgated as Financial Action Task Force (FATF) Recommendations is conducted by the IMF and World Bank Group, as well as FATF-style regional bodies associated with the FATF (FATF Secretariat, 2014). Currently, one notable monitoring and assessment mechanism in international taxation issues is the peer review process conducted by the OECD's Global Forum and its member jurisdictions. In order to monitor the implementation of the BEPS project globally, the *G20 Leaders' Communiqué* of the 2015 Antalya Summit 'call[ed] on the OECD to develop an inclusive framework' (G20, 2015) and, in response, the OECD launched a new inclusive framework for the implementation of the BEPS outputs in June 2016 (OECD, 2016c). Membership of the inclusive framework is open to all jurisdictions which commit themselves to the BEPS project, and its programme of work includes the development of a monitoring process on the implementation of the BEPS outputs focussed on four issues: addressing harmful tax practices (Action 4); preventing treaty abuse (Action 6); Country-by-Country Reporting (Action 13); and dispute resolution (Action 14) (OECD, 2016a). The OECD's *Background Brief* (2016) for the inclusive framework cites a preceding example of the Global Forum, which has been carrying out peer review on the implementation of the international standards on transparency and exchange of information. The *modi operandi* of the inclusive framework for monitoring the implementation of BEPS project's outputs will draw upon the experience of the Global Forum. In light of these developments, it appears that the Global Forum-style framework under the umbrella of the OECD is becoming a standard approach as a monitoring process in the international taxation field.

Lastly, 'Enablers', the new category added by Grinberg, are very important for international taxation in the G20 era. As stated in Section 3, the capacity of the OECD as a technical assistance provider is limited and yet, at the same time, a broad spectrum of implementation, including implementation in developing countries, is expected. The OECD therefore needs to strengthen its partnerships with technical assistance-savvy international organisations, such as the IMF and World Bank Group. In April 2016, the IMF, OECD, United Nations and World Bank Group jointly announced a plan to intensify their co-operation through a new Platform for Collaboration on Tax (the Platform), which evidences demand for co-operation amongst international organisations in assisting developing countries with respect to international taxation and other tax issues (IMF, OECD, UN & World Bank, 2016).

The concept note for the Platform indicates that its activities will also involve regional development banks, regional tax organisations and donors. Therefore, there are other organisations acting as Enablers (IMF, OECD, UN & World Bank, 2016). Regional development banks, such as the Asian Development Bank, and bilateral development agencies, such as Great Britain's Department of International Development and the German international

⁷ For example, in 2009, the Philippines were categorised by the OECD as not committed to information exchange standards and, in 2010, the Philippines enacted the Exchange of Information on Tax Matters Act of 2009, which enabled the Philippine Bureau of Internal Revenue to access to bank information for the purpose of exchanging information with foreign tax authorities.

⁸ According to *Tax Administration 2015*, an OECD's comparative analysis report, 32 of 56 jurisdictions surveyed have an autonomous tax administration body. Cf. OECD (2015c).

co-operation agency, GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH), are providing technical assistance for taxation issues. For example, in 2014, the Asian Development Bank launched a technical assistance project which aimed to enhance the capacity of tax administration bodies in Asia for the exchange of information and the investigation of cross-border tax evasion cases (Asian Development Bank, 2014).

Moreover, tax administration bodies are forming co-operative frameworks or forums on a regional basis; for example, there is the CIAT (Centro Interamericano de Administraciones Tributarias/Inter-American Center of Tax Administrations) in the Americas (Araki, 2015). These regional co-operation frameworks for tax administration bodies can also play a significant role in the capacity development of tax authorities in respective regions. The following sections will further discuss the importance of engagement with developing countries and the roles of regional co-operation frameworks.

5. ENGAGEMENT WITH DEVELOPING COUNTRIES IN THE G20 ERA

Section 2, on the current landscape, has argued that a paradigm shift in the international taxation field from the pre-G20 era to the G20 era in 2009 changed relationships with developing countries. This change has, in turn, influenced the composition of the architecture of international taxation standards, including the growing importance of ‘Enablers’, as pointed out in Section 4.

Before 2009, amongst the OECD’s international taxation agenda, engagement with non-OECD member jurisdictions was most active in respect of the transparency and information exchange issue, and yet focus was given to so-called tax havens and offshore financial centres.⁹ The OECD’s Global Forum on Taxation, the predecessor of the current Global Forum, started its work in 2000 with OECD member countries, and only six jurisdictions which had made political commitments to improve transparency and their information exchange systems (i.e. Bermuda, the Cayman Islands, Cyprus, Malta, Mauritius and San Marino) (OECD, 2006).

As described in Section 2, the G20 London Summit in April 2009 was a game changer (Grinberg, 2015, p.9). Whilst the primary target of the summit’s statement and declaration was offshore financial centres identified by the OECD’s work (London Summit – Leaders’ Statement, 2009), the declaration referred to the need for a more inclusive framework which could benefit ‘developing countries’ (Declaration on Strengthening the Financial System – London Summit, 2009). In September 2009, in response to the G20’s call, the OECD launched the restructured Global Forum. Whilst the restructured Global Forum opened its membership to all jurisdictions committed to implementing international standards on tax information exchange, its initial list of potential members included 91 jurisdictions consisting of: 30 OECD member countries; five candidate countries for OECD membership (including Russia); seven G20 member countries which were not OECD members; and 49 jurisdictions considered to be financial centres (OECD, 2009b, p5). In the meantime, the Global Forum recognised capacity building in developing countries as a key pillar of its activities, along with developing monitoring and peer review processes, and welcomed capacity building activities provided by other international and regional organisations (OECD, 2009a, p.3).

⁹ In 2000, the OECD listed 35 jurisdictions which met the OECD’s criteria for tax havens. Cf. OECD, Committee on Fiscal Affairs (2000), p.17.

As of January 2016, the Global Forum's members have risen to 129 jurisdictions and the European Union, and 38 countries have joined the forum in addition to the 91 jurisdictions listed in September 2009. These 38 countries are mostly developing nations, 16 of which are in Africa. The composition of the forum increased the weight of developing countries which have, in contrast to classic offshore financial centres, substantial population, economic activities and inland revenue. The September 2009 list included only four African countries, i.e. Liberia, Mauritius, Seychelles and South Africa, out of 91 jurisdictions (4 per cent), and yet, as at January 2016, the forum members include 20 African countries out of 129 jurisdictions (16 per cent).

In its early stages, the BEPS project recognised that challenges lay in not merely identifying solutions, but in implementing them in a streamlined manner (OECD, 2013a, p.8). The *Action Plan on Base Erosion and Profit Shifting*, released in July 2013, discussed engagement with developing countries as part of an inclusive and effective process (OECD, 2013b, pp.25-26). The OECD's reports, released in 2014, on the impact of BEPS in low income countries underlined the idea that the BEPS phenomena harmed both developed and developing countries, and that capacity development addressing the BEPS issues was critical for developing countries (OECD, 2014).

A number of developing countries have participated in the process of producing the BEPS project's outputs, via the Committee on Fiscal Affairs' meetings in Paris and regional network meetings. After the release of the BEPS project's outputs in October 2015, focus has shifted towards their implementation in a broad range of countries. Monitoring and technical assistance activities on the part of developing countries will carry considerable weight in the new inclusive framework for the implementation of the BEPS package, launched in June 2016, as touched upon in Section 4.

Work on international taxation is recognised as part of development agenda, particularly in the context of domestic resource mobilisation for development. The *Addis Ababa Action Agenda of the Third International Conference on Financing for Development*, which was endorsed by the United Nations General Assembly in July 2015, called for more inclusiveness to ensure that ongoing efforts in international tax co-operation (such as the work of the OECD's Global Forum and BEPS project) benefit all countries, including developing countries. The *Addis Ababa Action Agenda* also welcomed capacity-building and dialogue activities arranged by international organisations, such as the IMF, OECD and United Nations (United Nations General Assembly, 2015).

6. ROLES OF REGIONAL CO-OPERATION FRAMEWORKS

Drezner (2007), as introduced in Section 3, suggests 'neighbourhood international organisations' use geography to place a natural and fixed limitation on membership as a category of international governmental organisations; and as mentioned in Section 4, tax administration bodies are forming regional co-operation frameworks or forums. These frameworks can play significant roles in the architecture of international taxation standards, particularly in the capacity development of tax authorities, as well as bridging between 'Standard Setters' and respective regions.

Figure 3 : Major Regional Co-operation Frameworks for Tax Administration Bodies			
Organisation	Number of Members	Year of Establishment	Permanent Secretariat
Study Group on Asian Tax Administration and Research (SGATAR)	17	1970	None
Pacific Islands Tax Administrators Association (PITAA)	16	2003	Suva, Fiji
Inter-American Center of Tax Administrations (Centro Interamericano de Administraciones Tributarias/CIAT)	38	1967	Ciudad de Panamá
Commonwealth Association of Tax Administrators (CATA)	47	1978	London
Centre de rencontres et d'études des dirigeants des administrations fiscales (Meeting and Studies Centre of Tax Administration Directors /CREDAF)	30	1982	Paris
Intra-European Organisation of Tax Administrations (IOTA)	46	1996	Budapest
Association of Tax Authorities of Islamic Countries (ATAIC)	28	2003	Khartoum
African Tax Administration Forum (ATAF)	36	2009	Pretoria

Figure 3 lists major regional frameworks or forums for co-operation between the world's tax administration bodies (Araki, 2015). The membership of these frameworks is not merely based upon geography, but also on common cultural backgrounds. Whilst the organisational structures of these regional frameworks vary, they have similar mission statements (Alink & van Kommer 2016). Each framework aims to act as a platform for enhancing the capacity and performance of tax administration in a region, through international co-operation, the exchange of knowledge and experiences, and technical assistance activities.

As the focus of the BEPS project shifts from standard-making to implementation, particularly in a manner inclusive of developing countries, the roles of regional frameworks for co-operation have been realised anew. The *G20 Response to 2014 Reports on Base Erosion and Profit Shifting and Automatic Exchange of Tax Information for Developing Economies*, which is an annex prepared by the G20 Development Working Group for the 2014 Brisbane G20 Summit in November 2014, 'recognise[d] the importance of regional (including inter-regional) tax administration forums in creating a bridge between the international tax agenda and developing economies'.

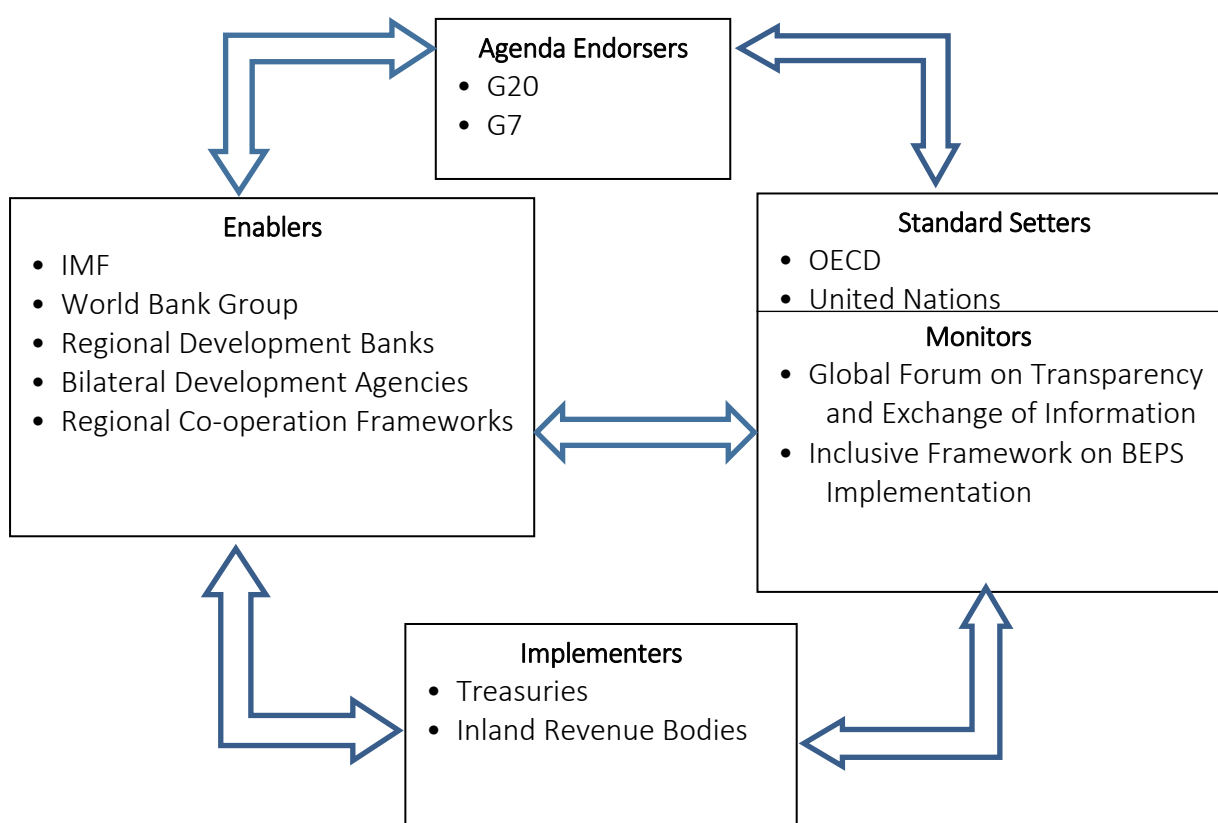
Indeed, regional co-operation frameworks are already taking substantial roles in the process of making and implementing international standards in international taxation. For example, the Centro Interamericano de Administraciones Tributarias/Inter-American Center of Tax

Administrations (CIAT) and African Tax Administration Forum (ATAF) are attending discussions at the OECD's Committee on Fiscal Affairs (CFA) as observers along with the IMF, United Nations and World Bank Group.¹⁰ As mentioned in Section 5, the OECD has organised a series of regional networking meetings in order to engage with developing countries in the BEPS project, and many, if not all, of these regional networking meetings have been held in partnership with regional co-operation frameworks, notably the CIAT, ATAF, Centre de rencontres et d'études des dirigeants des administrations fiscales (Meeting and Studies Centre of Tax Administration Directors /CREDAF) and Intra-European Organisation of Tax Administrations (IOTA).

7. AN ARCHITECTURE MODEL OF INTERNATIONAL TAXATION STANDARDS

In light of discussions about the players of the international taxation field in the sections above, Grinberg's model on the architecture of international financial law, presented as Figure 2, in Section 2, may be further modified as follows:

Figure 4: A Modified Architecture Model of International Taxation Standards



¹⁰ Source: On-Line Guide to OECD Intergovernmental Activity, Committee on Fiscal Affairs. Retrieved from <http://webnet.oecd.org/OECDGROUPS/Bodies/ListByNameView.aspx?book=true>

Brummer (2015) points out institutional interdependency as one of the characteristics of the international regulatory system.¹¹ The architecture of international taxation standards can be thought of as a process in which several players grouped by each function mutually interact, and expressed by bidirectional arrows, rather than a top-down process centred round the G20. As argued in Section 4, although the G20 is the premier international forum, the kernel of its role can be seen as, if anything, giving political legitimacy as global agenda to projects undertaken by the OECD as a 'Standard Setter', the official membership of which is inclined to developed countries.¹²

In the case of international taxation standards, the monitoring function is provided by frameworks arranged by the OECD with open membership; the Global Forum on Transparency and Exchange of Information for Tax Purposes is a prime example. As argued in Section 4, it is probable that a new inclusive framework for the implementation of the BEPS project's outputs, launched in June 2016, will take a similar route.

'Enablers' have an extended list of functions. The *G20 Leaders' Communiqué* of the 2015 Antalya Summit welcomed the efforts made by the IMF and World Bank Group to provide technical assistance to developing economies in order to help them tackle domestic resource mobilisation challenges such as international taxation issues (G20 Leaders' Communiqué, 2015). As discussed in Section 4, in April 2016, the IMF, OECD, United Nations and World Bank Group jointly announced a plan to strengthen their co-operation through a new Platform for Collaboration on Tax, and its objectives include fostering more dynamic interactions between 'Standard Setters' and technical assistance providers (IMF, OECD, UN & World Bank, 2016). As mentioned in Section 6, the IMF, UN and World Bank Group are also attending the OECD's Committee on Fiscal Affairs (CFA) as observers.

Other regional development banks and bilateral development agencies are also conducting research and technical assistance activities, mostly in a regional context. Section 6 has argued that regional co-operation frameworks are playing a significant part in the capacity development of their members' tax administration bodies.

Whilst 'Implementers' are put at the bottom of the architecture, international standards do not take effect until they are implemented; in other words, implementation is the primary output of the architecture of international taxation standards. In addition, the architecture can be as bottom-up, just as it can be top-down. It is fair to say that reforms in international taxation rules and standards are demand-driven, and issues are identified and discussed by 'Implementers' *per se* on the stage set by 'Standard Setters'.¹³ As mentioned in Section 2, in addition to 34 official member countries of the OECD, other countries (including emerging G20 members) have directly joined discussions at the OECD's CFA with regard to the BEPS project.

¹¹ Brummer points out that the G20 has no permanent staff, and relies on 'Standard Setters', such as the Basel Committee on Banking Supervision, to develop standards and to provide input on regulatory direction; and in a similar vein, these 'Standard Setters' also rely on monitors to track national regulations (pp. 73 and 116).

¹² Brummer also argues that a degree of participation, particularly of developing countries, in the rulemaking process raises the question of legitimacy, which, at least in part, spurred the recent displacement of the G7 by the G20, and the expansion of membership in standard setting processes (pp.114-115).

¹³ In May 2014, at a dinner speech given for a G20 International Tax Symposium held in Tōkyō, the director of the OECD's Centre for Tax Policy and Administration recalled that the BEPS project had been conceived as a result of discussion amongst senior tax officials, who were members of the Committee on Fiscal Affairs, on issues with which tax authorities were faced.

Furthermore, the largest ‘Implementers’ - governments - comprise the G20,¹⁴ and are major shareholders in international financial institutions as well.¹⁵

8. CONCLUSION

Although international taxation as an international issue dates back to the nineteenth century and multilateral work started with the League of Nations, as introduced in Section 1, interest in the global framework or architecture of international taxation issues as a transgovernmental system was not as great as that in other financial initiatives, such as those undertaken by the Basel Committee on Banking Supervision and Financial Action Task Force (FATF).

As shown in Section 2, interest in international taxation issues as a global agenda started to grow after about 2009, with the rise of the G20 and the launch of the restructured Global Forum (cf. Porter & Rubio Vega, 2011). It has attracted further attention since June 2011, following the launch of the BEPS project, which aimed at a comprehensive reform of the international taxation system.

Implementation is no less important in taxation issues than in other financial issues, and its significance cannot be exaggerated. In this regard, as the *G20 Leaders' Communiqué* of the 2015 Antalya Summit made clear, ‘widespread and consistent implementation will be critical in the effectiveness of the [BEPS] project’ (G20 Leaders' Communiqué, 2015). And the thing that ensures that implementation is widespread is engagement with developing countries. In the pre-G20 era, international taxation standards were chiefly set as a procedure within the OECD (Grinberg, 2015). In contrast, in the G20 era, a global framework which aims to cover the implementation of international standards in developing countries within its range is required to be extended as a network of ‘players’ in the international taxation field.

The global architecture or framework of international taxation standards is not yet fully fledged, when compared to other financial initiatives, and it is still in the process of evolving. There is no doubt that the primary focus is to make the global framework inclusive of emerging and developing countries. On the other hand, as pointed out in Sections 3 and 4, the capacity of the OECD and its frameworks as a technical assistance provider or ‘Enabler’ is not sufficient. A second focus for the global framework is to establish functional partnerships which enable international organisations to address issues faced by developing countries and other ‘Implementers’ effectively. Lastly, in order for the architecture to have a truly global reach, it will be required to establish a solid network through regional frameworks or forums.

¹⁴ The G20 membership is said to represent about two-thirds of the world’s population and 85 per cent of global gross domestic product. Source: G20 2015 Turkey. G20 Members. Retrieved from <http://g20.org.tr/about-g20/g20-members/>

¹⁵ For example, the G20 members have approximately 65 per cent of voting shares in the IMF. Source: IMF. IMF Members' Quotas and Voting Power, and IMF Board of Governors. Retrieved from <https://www.imf.org/external/np/sec/memdir/members.aspx>

REFERENCES

- Alink, M. & van Kommer, V. (2016). *Handbook on Tax Administration* (Second Revised Edition). Amsterdam: IBFD.
- Araki, S. (2015). Regional Cooperation and Tax Information Exchange among Asia-Pacific Tax Authorities. *Asia-Pacific Tax Bulletin*, 21. Amsterdam: IBFD.
- Asian Development Bank (2014). *Technical Assistance Report: Enhancing Transparency and Exchange of Information for Tax Purposes*. Metro Manila. Retrieved from <http://www.adb.org/sites/default/files/project-document/80788/47149-001-tar.pdf>
- Ault, H. (2009). Reflections on the Role of the OECD in Developing International Tax Norms. *Brooklyn Journal of International Law*, 34, 760-761.
- Brummer, C. (2015). *Soft Law and the Global Financial System: Rule Making in the 21st Century* (Second Edition). New York: Cambridge University Press.
- Drezner, D. (2007). *All Politics Is Global*. Princeton: Princeton University Press.
- FATF Secretariat. (2014). *Financial Action Task Force: Annual Report 2013-2014*. Paris: FATF/OECD.
- Grinberg, I. (2015). Breaking BEPS: The New International Tax Diplomacy. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2652894
- Global Forum on Transparency and Exchange of Information for Tax Purposes. (2009). *Tax Co-operation 2009: Towards a Level Playing Field*. Paris: OECD.
- London Summit – Leaders’ Statement (2009).
- Declaration on Strengthening the Financial System – London Summit. (2009).
- G20 Leaders Statement: The Pittsburgh Summit. (2009).
- G20 Leaders Declaration, G20 Los Cabos Summit. (2012).
- G20 Leaders' Communiqué, G20 Antalya Summit (2015).
- G7 Ise-Shima Leaders’ Declaration, G7 Ise-Shima Summit. (2016).
- Haas, P. (1992). Introduction: Epistemic Communities and International Policy Coordination. *International Organization*, 46(1), 3.
- IMF, OECD, UN & World Bank. (2016). *The Platform for Collaboration on Tax: Concept Note*.
- Jogarajan, S. (2011). Prelude to the International Tax Treaty Network: 1815–1914 Early Tax Treaties and the Conditions for Action. *Oxford Journal of Legal Studies*, 31(4), 679-707.
- OECD, Committee on Fiscal Affairs. (2000). *Towards Global Tax Co-operation: Progress in Identifying and Eliminating Harmful Tax Practices*. Paris.
- OECD, Global Forum on Taxation. (2006). *Tax Co-operation: Towards a Level Playing Field*. Paris.
- OECD. (2009a) *Summary of Outcomes of the Meeting of the Global Forum on Transparency and Exchange of Information for Tax Purposes Held in Mexico on 1-2 September 2009*.
- OECD. (2009b). *Decision of the Council Establishing the Global Forum on Transparency and Exchange of Information for Tax Purposes*. Paris.
- OECD. (2013a). *Addressing Base Erosion and Profit Shifting*. Paris.

- OECD. (2013b). *Action Plan on Base Erosion and Profit Shifting*. Paris.
- OECD. (2014). *Two-Part Report to G20 Developing Working Group on the Impact of BEPS in Low Income Countries*. Paris
- OECD. (2015a). *Explanatory Statement, OECD/G20 Base Erosion and Profit Shifting Project*. Paris.
- OECD. (2015b). *Mandatory Disclosure Rules, Action 12 – 2015 Final Report, OECD/G20 Base Erosion and Profit Shifting Project*. Paris.
- OECD. (2015c). *Tax Administration 2015: Comparative Information on OECD and Other Advanced and Emerging Economies*. Paris.
- OECD. (2016a). *Background Brief: Inclusive Framework for BEPS Implementation*. Paris.
- OECD. (2016b). *Statement from OECD Secretary-General Angel Gurría on the “Panama Papers”*. Retrieved from <http://www.oecd.org/tax/statement-from-oecd-secretary-general-angel-gurria-on-the-panama-papers.htm>
- OECD. (2016c). *First meeting of the new inclusive framework to tackle Base Erosion and Profit Shifting marks a new era in international tax co-operation*. Retrieved from <http://www.oecd.org/ctp/first-meeting-of-the-new-inclusive-framework-to-tackle-base-erosion-and-profit-shifting-marks-a-new-era-in-international-tax-co-operation.htm>
- Porter, T. & Rubio Vega, V. (2011). Global Forum on Transparency and Exchange of Information for Tax Purposes. In T. Hale & D. Held (Eds.), *Handbook of Transnational Governance: Institutions and Innovations*. Cambridge: Polity Press.
- Russo, R. (2016). 3 BEPS Years, a (very) personal perspective. Available on LinkedIn.
- Sarmiento, D. (2010). The Function of EU Soft Law. In D. Weber (Ed.), *Traditional and Alternative Routes to European Tax Integration*. Amsterdam: IBFD.
- Schrauwen, A. (2010). Sources of EU Law for Integration in Taxation. In D. Weber (Ed.), *Traditional and Alternative Routes to European Tax Integration*. Amsterdam: IBFD.
- Slaughter, A. (2004). *A New World Order*. Princeton: Princeton University Press.
- United Nations, Department of Economic & Social Affairs. (2001). *United Nations Model Double Taxation Convention between Developed and Developing Countries*. New York.
- United Nations, Economic and Social Council. (2004). *ECOSOC Resolution 2004/69: Committee of Experts on International Cooperation in Tax Matters*.
- United Nations, General Assembly. (2015). *Resolution adopted by the General Assembly on 27 July 2015: 69/313. Addis Ababa Action Agenda of the Third International Conference on Financing for Development (Addis Ababa Action Agenda)*. New York.

COMMENT ON FEIGE'S PAPER "REFLECTIONS ON THE MEANING AND MEASUREMENT OF UNOBSERVED ECONOMIES: WHAT DO WE REALLY KNOW ABOUT THE 'SHADOW ECONOMY'?"¹

Friedrich Schneider²

Abstract

This comment provides a reply to Prof. Feige's paper with the title "Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the 'Shadow Economy'?", in which Prof. Feige heavily criticizes me. I show that the same critique which Prof. Feige raises against me can be put forward to his results on the non-observed economy. Moreover, I show that my dataset is appropriately documented, and I also address the problem of calibration and normalization issues when undertaking a MIMIC estimation. In the concluding section, I suggest that a joint paper should be written in which all the pros and cons of each method of estimating the size and development of the non-observed/shadow economy are presented and criticized.

1. INTRODUCTION

In the paper "Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the 'Shadow Economy'?", Prof. Edgar L. Feige scientifically attacks me. He writes in the abstract: "... It concludes that Schneider's Shadow Economy (SSE) estimates suffer from conceptual flaws, apparent manipulation of results and insufficient documentation for replication, questioning their place in the academic, policy and popular literature."³ In this comment, I will react to this severe, heavy and, in my opinion, quite unfair criticism.

In Section 2, I will make some remarks about Prof. Feige's views on measuring the non-observed economy (NOE). Section 3 handles Prof. Feige's claim of insufficient documentation for replication. In Section 4, I will deal with the problem of calibration and normalization issues when undertaking a MIMIC estimation. In Section 5, the concluding section, I will make some final remarks about the MIMIC estimation procedure and other estimation procedures.

¹ Forthcoming in the Journal of Tax Administration, 1/2016, I would like to thank the editors, Prof. Dr. Chris Heady, Nigar Hashimzade and G. D. Myles for their kindness in allowing me to write this comment. The author is indebted to, and thanks gratefully, his co-authors (Andreas Buehn, Roberto Dell'Anno, Dominik Enste, Lars Feld) for the stimulating comments and help in writing this comment.

² Prof. Dr. Friedrich Schneider, Department of Economics, Johannes Kepler University of Linz, Altenberger Strasse 69, A-4040 Linz-Auhof, Austria. Phone: +43-732-2468-7340, Fax: +43-732-2468-7341, E-mail: friedrich.schneider@jku.at, www.econ.jku.at/Schneider

³ I would like to mention that I have never criticized Prof. Feige in such a severe and drastic way. I always quote him in my papers because he invented the transaction approach to measuring the size and development of the shadow economy and has made other contributions in the debate about measuring the size of the shadow economy. I am aware of his achievements and, hence, have always been quite fair about him in my papers. Compare e.g. Schneider and Enste (2000), Schneider, Buehn and Montenegro (2010) and Schneider and Williams (2013).

2. FEIGE'S PRESENTATION OF MEASURING THE NON-OBSERVED ECONOMY/INCOME

In his paper, Prof. Feige heavily relies on the following publication: OECD (2002), "Measuring the non-observed economy: A handbook" (Paris: France). In his Table 1, he presents results about the size and development of the measured non-observed income (Y_{NOE}^m) or measured non-observed economy of 15 former Soviet Union countries (from Armenia to Uzbekistan) and of 12 Central Eastern and Eastern European countries (from Albania to Slovenia) using a discrepancy approach. This is an interesting table with quite challenging results. I will make the following critical remarks:

2.1 The Sources of Table 1: Can One Replicate the Results?

For the 27 countries, the sources of the reported values of the measured non-observed economy (Y_{NOE}^m) are personal interviews and personal correspondence, including some with national statistical offices. Here Prof. Feige should first explain:

- What does this mean?
- Were these figures transmitted to Ed Feige by personal correspondence?
- Is this an official source for this country or not?
- How can one examine how these figures are calculated using only this source?

At the very least, much more careful documentation should be provided here, so that everyone can check how reliable these results are. This is not possible from the sources given in Table 1.

2.2 Accuracy of the Shadow Economy Figures

On page 20 in his paper and in the part "Schneider's Shadow Economy (SSE)", Ed Feige attacked me: "...to have estimated the size and trend of the shadow economy worldwide, for 162 countries to the accuracy of one decimal place...". I am somewhat amazed, because in Tables 1 and 2, Ed Feige produces results of the size of different non-observed economies in exactly the same way (e.g. for Armenia, for the year 1995, of 31.6% of GDP). However, I always state that these point estimates have a margin error of $\pm 15\%$, while he does not say anything about the possible error of his figures.

2.3 Drastic Changes in the Size of the Shadow Economy

If I look at his figures, there are remarkable changes. Russia had a non-observed economy of 11.9% in 1997 and 1998 then, all of a sudden, in 2000, it was 24.8% and, in 2003, 24.3%; so the size of the non-observed economy doubled over two years. In Bulgaria, the size of the non-observed economy was 27.8% in 1996 and 31.2% in 1997, and fell to 12.3% and 12.0% in the years 1998 and 1999. Hence, in the years 1998 and 1999, the size of non-observed economy was only about a third of the previous year's. The value of the non-observed economy of Hungary was always "precisely" 16.0% for the years 1992 to 1999 (exactly the same value for 8 years). Then, in the year 2000, it dropped to 11.9% and, in 2001, it was again 16.0%. How can this be? No explanation is given. The non-observed economy in Slovenia was between 6.4% and 7.5% from 1995 to 2002, then suddenly keeps stable at "precisely" 8.0% from 2003 to 2007 (5 years). I really think all these changes need explanations, especially if Prof. Feige

heavily criticizes the size of the non-observed economies of my co-authors and me, even though we never have such jumps in the figures.

2.4 Different Sizes of the Shadow Economy: The Case of Serbia

To give an example as to how the size of these figures might also cause severe doubts about their reliability, I take the example of Serbia. In Krstić and Schneider (2015), we tried to measure the size and development of the shadow economy in Serbia using various, methodologically different, approaches. The MIMIC procedure was only one of them. We also used the household tax compliance (HTC) method and the survey method, and got the following results:

- The estimated value of the shadow economy, according to the MIMIC method, in the year 2010 was 30.1% of GDP;
- according to the HTC method, it was 23.6% in 2010 and;
- according to the survey method, it was 21.0% of GDP in 2012.

Hence, even with the other methods, which have nothing in common with the MIMIC approach (especially the survey method), much higher values of the shadow economy are estimated, compared to Ed Feige's result of 14.6% in the year 2003.

2.5 The Difference of the Size of the Shadow Economy Between Macro and Other Estimates

Finally, I want to raise the point of whether the size and development of the shadow economy estimated by using a macro approach (either the MIMIC or a currency demand approach) can be so easily compared to the estimates of the non-observed income method which Feige uses. One disadvantage of the macro methods⁴ is that estimations quite often include legally-bought material when showing the estimated size and development of the shadow economy. If one employs a shadow economy worker, he goes to the construction market and buys the necessary things which the hired shadow economy worker needs to do his job, but these are legally-bought goods which are already counted in official GDP and which are taxed. In order to compare macro approaches with other approaches, one should subtract the legally-bought material.⁵ If one undertakes such a correction, one can make the assumption that roughly 20% needs to be deducted from these macro shadow economy measures to allow for legally-bought material which is already counted in official GDP.

In Table 1, an example for the size of the German shadow economy is given in order to explain the estimated difference between the survey approach, which traditionally results in much lower estimated values, and the MIMIC approach. Table 1 clearly shows that the MIMIC approach for the German shadow economy reached results between 15.5% and 16.0% of GDP (100% normalized), while material used accounts for 19.0% to 25.0% of the total shadow economy, illegal activities account for 27.0% to 30.0%, and shadow economy activities already included in official GDP make up 6.0% to 12.0%. Table 1 nicely demonstrates that we have

⁴ I have raised this question in several publications, e.g. in Enste and Schneider (2006), Schneider and Williams (2013), and Williams and Schneider (2016).

⁵ This is a difficult task, as we do not have good data on this. Data exists only for Germany, where several surveys have been undertaken to use a micro approach for estimating the shadow economy. Compare the references in Table 1.

different results with respect to the method used to estimate the size and development of the shadow economy.

Table 1: Size of the shadow economy in Germany in the year 2005 using different estimation approaches

Estimation approach	in % of off. GDP	in Bill. Euros	in % of the total shadow economy in Germany
Survey about black labor as value-added provided by Feld and Larsen (2012a)	3.6%	70	22.5%
+ corrections of the survey, see e.g. Feld and Larsen (2012a, p. 61)	5.1%	112	32%
+ material used	3.0–4.0%	65–90	19–25%
+ illegal activities	4.3–4.8%	90–105	27–30%
+ already in the official GDP included shadow economy activities	1.0–2.0%	20–40	6–12%
Shadow economy using the MIMIC procedure (for calibration the currency demand approach)	15.5–16.0%	340–350	100%

Table source: Enste and Schneider (2006), Table 2, p. 188.

Sources of the representative survey: Feld and Larsen (2005, 2012a, b) and Pedersen (2003).

The source of illegal activities and official material used are based on a survey of TNS-Emnid (2004) ordered by the German research institute IW, Cologne.

In his Table 2, Ed Feige undertakes a comparison with my estimates for the former Soviet Union countries, the Central Eastern European countries and some Western OECD countries. He took the averages from his Table 1, with widely different time series for the 27 countries (from one available year up to 12 available years), and from a study of the United Nations Economics Commission for Europe (2008). What is really amazing is that the Netherlands should only have, according to the accounting method, a shadow economy of 1%, Norway of 1.7%, Sweden of 1.3% and Turkey of 1.7%; remarkably low values. With the same measurement method, Austria has a shadow economy of 7.9%, which is remarkably high. No explanation is given. Just the values are taken. If I make a comparison over the years 2000 to 2002, for which Feige has values for most countries, and compare them with my figures in Schneider, Buehn and Montenegro (2010), I get the results shown in Table 2 (below). Moreover, in Table 3, I show a comparison between the UN estimates (UN, 2008) which Ed Feige used, and my estimates for OECD and European Union members. In both Tables 2 and 3, I also deducted 20% in one column for used material (officially bought) and one sees that, even without the correction, the difference is not as large as Ed Feige reports in his Table 2.⁶ Feige's Table 2 leaves at least two questions open:

⁶ The extremely low values of the Netherlands with 1.0% and Sweden with 1.3% are not plausible at all. Compare e.g. Kazemier (2006) and Williams and Schneider (2016, p. 53), where Kazemier reports a result of 9.1% using the survey method.

- (i) What is the precise period of the average? Is this period different from country to country? If so, it should be stated.
- (ii) How can an interested researcher verify these results and from where does he get the necessary documentation?

Table 2: A comparison of Feige's and Schneider's (with co-authors) average estimates for the non-observed economy over 2000 to 2002 in % of GDP

Country	Feige's estimates				Schneider's estimates					Ratio S.Av./ F.Av.	Ratio reduced S.Av./ F.Av.
	2000	2001	2002	Av.	2000	2001	2002	Av.	Av. (- 20%)		
Albania	34.2	30.4	30.5	31.7	35.3	34.9	34.7	35.0	28.0	1.10	0.88
Armenia	30.2	28.2	29.4	29.3	46.3	45.4	44.5	45.4	36.3	1.55	1.24
Azerbaijan	19.5	22.7	19.2	20.5	60.6	60.3	60.0	60.3	48.2	2.95	2.35
Belarus	11.1	10.6	11.1	10.9	48.1	47.9	47.6	47.9	38.3	4.38	3.52
Bulgaria	16.3	10.2	n.v.	13.3	36.9	36.6	36.1	36.5	29.2	2.76	2.20
Croatia	8.5	8.3	8.2	8.3	33.4	33.2	32.6	33.1	26.5	3.97	3.19
Czech Rep.	7.7	7.5	6.9	7.4	19.1	18.9	18.8	18.9	15.1	2.57	2.04
Estonia	8.9	7.4	9.6	8.6	32.7	32.4	32.0	32.4	25.9	3.75	3.01
Georgia	33.7	33.4	33.2	33.4	67.3	67.2	67.2	67.2	53.8	2.01	1.61
Hungary	11.9	16.0	n.v.	14.0	25.1	24.8	24.5	24.8	19.8	1.78	1.42
Kazakhstan	24.7	23.9	22.6	23.7	43.2	42.5	42.0	42.6	34.1	1.79	1.44
Kyrgyzstan	13.1	14.4	16.5	14.7	41.2	40.8	41.4	41.1	32.9	2.80	2.24
Latvia	18.0	17.5	16.0	17.2	30.5	30.1	29.8	30.1	24.1	1.76	1.40
Lithuania	18.0	18.3	18.9	18.4	33.7	33.3	32.8	33.3	26.6	1.81	1.45
Macedonia	12.9	14.9	14.4	14.1	38.2	39.1	38.9	38.7	31.0	2.75	2.20
Moldova	34.6	31.6	n.v.	33.1	45.1	44.1	44.5	44.6	35.7	1.35	1.08
Poland	14.6	14.3	15.4	14.8	27.6	27.7	27.7	27.7	22.2	1.87	1.50
Romania	21.1	n.v.	17.7	19.4	34.4	33.7	33.5	33.9	27.1	1.75	1.40
Russia	24.8	n.v.	n.v.	24.8	46.1	45.3	44.5	45.3	36.2	1.83	1.46
Slovakia	14.9	15.2	14.6	14.9	18.9	18.8	18.6	18.8	15.0	1.26	1.01
Slovenia	6.9	6.8	7.5	7.1	27.1	26.7	26.6	26.8	21.4	3.79	3.02
Ukraine	20.0	16.3	17.7	18.0	52.2	51.4	50.8	51.5	41.2	2.86	2.29

n.v. = no value; Sources: Schneider, Buehn and Montenegro (2010, pp. 454–456); Feige (2016, p. 14).

Table 3: A comparison of UN (used by Feige) and Schneider's (with co-authors) estimates for the non-observed economy for some OECD-EU members for various years in % of GDP

Country	Year	UN estimates	Schneider's estimates	Schneider's estimates reduced by 20%	Ratio Schneider/UN	Ratio Schneider (reduced)/UN
Austria	2001	7.9	9.7	7.8	1.23	0.98
Belgium	2002	3.0–4.0	22.0	17.6	5.50–7.33	4.40–5.87
Finland	2001	Not stated	17.9	14.3	-	-
Germany	2001	Not stated	15.9	12.7	-	-
Ireland	1998	4.0	16.1 (1999)	12.9	4.03	3.22
Italy	2003	14.8–16.7	27.0	21.6	1.62–1.82	1.29–1.46
Netherlands	1995	1.0	13.3 (1999)	10.6	13.30	10.64
Spain	2000	11.2	22.7	18.2	2.03	1.62
Sweden	2000	1.3	19.2	15.4	14.77	11.82
United Kingdom	2001	Not stated	12.6	10.1	-	-

Sources: Schneider, Buehn and Montenegro (2010, pp. 454–456); UN (2008, p.10).

3. INSUFFICIENT DOCUMENTATION OF THE DATA USED BY MY CO-AUTHORS AND ME

Another heavy claim and criticism in Ed Feige's paper is, as Ed Feige writes: ... "It concludes that SSE estimates suffer from ... insufficient documentation for replication...". This is a very strong claim which is absolutely not true. My co-authors (Andreas Buehn, Roberto Dell'Anno, Egle Tafenau, Helmut Herwartz, Dominik Enste, Lars Feld etc.) and I always take great care that everyone who wants the dataset can have it. The most requested dataset is the one from Schneider, Buehn and Montenegro (2010), where estimates for 162 countries were published for the period 1999 to 2006/07 in the *International Economic Journal*, Vol. 24/4. I have sent this dataset to so many interested researchers that I have stopped counting them and, let me clearly say, Ed Feige also received access to this dataset. In a mail from July 16, 2012, 9:30 pm, Andreas Buehn sent Ed Feige detailed information, including a description of calibration methodology and five papers that provide background information on the MIMIC procedure itself, different calibration methods as well as exogenous estimates used for calibration. In a mail to Ed Feige from July 17, 2013, 12:27 pm, I sent Ed Feige requested documentation, LISREL program code, and an explanation and description of the variables for the World Bank paper again. My co-authors and I gave him several specifications and we provided him with the description of variables used in the model, the definition of the variables used in the model, and detailed documentation relating to the World Bank paper, including the list of variables and the do-files. He also got the data which we used in the World Bank working paper and later on in the published paper.⁷ Moreover, on June 7, 2013, I sent him documentation relating to the size of the shadow economies in 179 countries, in which the figures for the MIMIC calibration procedure, starting in Albania and ending with Zimbabwe, were shown. I provided

⁷ Andreas Buehn did send him the data in a zip-file on May 16, 2012.

him the exact sources; for example, for Albania, there were 7 sources, where 3 sources were from quite different authors. Later on, I sent him another document with the title "A preliminary documentation of the size of the shadow economy of 27 selected countries: The figures for the MIMIC calibration procedure", which completed my first dataset and, in this second piece of documentation, none of the sources are from me or my co-authors. I really think one cannot do more. In a mail from November 25, 2013, I asked him whether he had received all the necessary data and in the third point of my mail I asked him: "Would you be willing to send me a similar detailed documentation of the econometric estimates of your last paper, I mean the paper Feige & Urban published in 2008, in order to estimate the size and development of the shadow economy?". I never got an answer; I never got the data, not even a reply.

To summarize, I have provided Ed Feige with all the necessary data.⁸ I believe I have fulfilled all tasks, and Ed Feige never reacted to my requests to send me his data and do-files, in order to be better able to understand what he did in his paper "Measuring Underground (Unobserved, Non-Observed, Unrecorded) Economies in Transition Countries: Can we Trust GDP?", jointly written with Ivica Urban. Hence, I really think his statement is not true at all.⁹

4. CALIBRATION AND NORMALIZATION ISSUES WHEN UNDERTAKING A MIMIC ESTIMATION

Here I want to raise two points. The first is which of the indicator variables should be normalized and, further, whether it makes sense to normalize GDP with -1 . In the papers of my co-authors (Dell'Anno, Buehn, Enste, Herwartz, Tafenu, Feld) and myself, we normally assume a coefficient of real GDP of -1 (in 12 out of 13 cases, see Table 3 of Ed Feige), which is derived from the theory that an increase in shadow economy activities has a negative effect on official GDP development. An increase in the shadow economy absorbs labor resources from the official economy, reducing labor supply in the official economy. Hence, assuming a negative coefficient is theoretically absolutely plausible. It is also absolutely plausible to assume that the coefficient of variable measuring currency holdings, if such an indicator variable was used, has a value of $+1$, because the higher the shadow economy, the higher currency holdings should be. Also it is theoretically plausible to assume that the average official working time per week can be normalized to -1 , as the more people work in the official economy, the less time they have to work in the shadow economy. Hence, these assumptions are not arbitrary, but based on theoretical grounds.

To summarize: the logic behind choosing the reference indicator GDP and its associated sign of -1 is the reasoning that the shadow economy absorbs human capital and resources from the official economy, leading to negative effects. It is also theoretically highly plausible that the higher the tax burden, the more regulation and the lower the tax morale, *ceteris paribus*, the higher the shadow economy will be. Most studies show this (compare Schneider and Enste, 2000, and Schneider, 2015). Hence, this is not an arbitrary choice! Moreover, instead of normalizing GDP to -1 , the variable "currency holdings" is often used and normalized to $+1$, assuming that the higher the shadow economy, the higher the amount of cash used, *ceteris paribus*. This alternative produces the same results as a positive influence of the tax rate, of the regulation index and a negative influence of tax morale, and these results are completely independent from the normalization of GDP.

⁸ Ed Feige admits, for example, that Breusch (2005) succeeded in replicating the earlier study of Dell'Anno and Schneider (2003) and the Asia-Pacific study by Bajada and Schneider (2005).

⁹ What is also not true is that we provided the data in 2016, which he writes in footnote 34 in his paper. We provided the data in 2012 and 2013, hence, over three years ago.

The second point that Prof. Feige discussed at length was a mistake that Andreas Buehn and I made when undertaking the calibration of the shadow economy values for 162 countries. Unintentionally, a sign error occurred in an Excel file. Unemployment was shrinking for almost all countries during the years 1999 to 2007 and, due to the mistake, the positive coefficient of unemployment was multiplied by -1 . Hence, we found an increase in the shadow economy. Unfortunately, this mistake, which we did not notice immediately, occurred. Realizing the mistake, we immediately corrected it, updated the dataset and published a revised version of the (working) paper. Such calculation errors can happen and I think we were completely right to correct our dataset, because we have a positive coefficient which is statistically significant in specifications 3, 4, 5 and 6. It was only not statistically significant in specifications 1, 2 and 7 (compare Table 1 in Schneider, Buehn and Montenegro (2010), p. 449).

Again, let me summarize: Andreas Buehn and I regret the calibration error but such mistakes can happen. We have corrected it and published a revised version of our work. Most importantly, it is absolutely not true that we did not offer a further explanation, admitting our mistake – it was an error in the calibration process and not manipulation of the dataset.

5. CONCLUDING REMARKS AND WHAT CAN WE LEARN?

In this reply, I have tried to demonstrate that the quite strong and extreme accusations Ed Feige makes against the estimates of the size and development of the shadow economy made by my co-authors and myself are really not justified. Of course, we all have to learn and we are aware that the MIMIC estimation procedure is a difficult one; one has to be careful when applying it and must also point out its weaknesses. In many of my contributions, I have done this. However, all methods trying to capture the size and development of the shadow economy should be treated in the same way. In addition, the discrepancy approach between national income and income statistics, and similar, related approaches have also been criticized (see, for example, Schneider and Enste, 2000, and especially Thomas, 1992).

Let me put forward another type of argument: I do not find criticizing the currency demand approach and the MIMIC approach in such a drastic way that these two approaches cannot and should never be used a scientifically useful and stimulating method of achieving progress in economics. This is more or less what Ed Feige is arguing. If I consider the methods he uses, especially the national income accounting framework and, deriving from this, the size of non-observed income and/or the non-observed economy, it is obvious that this approach, which is most often used by national accountants, is as problematic as the currency demand and/or MIMIC method, especially as it is very difficult to verify the sizes of the shadow economies calculated by this approach. Except for national accountants in national statistical offices, no one has access to the data and no one can see what the crucial assumptions are and how the figures were precisely calculated. This is also not done by Ed Feige in this paper. Hence, we have the problem that the results of these approaches are difficult, if not impossible, to replicate, and that we do not know assumptions and therefore do not know the variety of results. If one analyzes the results in Table 1 by Ed Feige, one observes huge jumps in the estimated values, e.g. a doubling of the non-observed economy or a shrinking of the non-observed economy by 50% within a year. All is possible.¹⁰ I really think one should criticize all possible methods to

¹⁰ Ed Feige does not provide the necessary sources and documentation to make it possible to check and verify the results in Table 1 and Table 2 in his paper. Neither the measurement of non-observed income by national statistical accounts in Feige's paper (2016), nor Tables 4 and 5 in the paper by Feige and Urban (2008), can be verified.

estimate the non-observed and/or shadow economy, and not just pick two of these methods and claim that they are not applicable because a lot of possible errors may happen. The MIMIC and currency demand approaches have the big advantage that everyone can use the data, make his or her own calculations, and compare them to the existing results both from these two macro methods and other methods. This is not possible with approaches which rely on national income statistics. To conclude, I am convinced that in order to estimate the size and development of a non-observed or shadow economy one should use all possible approaches, carefully explain the advantages and disadvantages, and compare the results. Then one might be able to come closer to a realistic value for the size of a shadow or non-observed economy. Furthermore, micro studies investigating why people work in the shadow economy and how much they work on an aggregate level are an interesting complement to the macro approaches, as I showed, for example, for the cases of Serbia and Germany. I think we should stop condemning the two macro approaches, be similarly critical towards the other ones, and compare the results of all approaches. Only then will we make scientific progress.

Let me conclude by saying that I really regret this dispute. I think it would be much more productive if Prof. Feige and I were to write a joint paper, clearly pointing to the differences between the various methods, but showing the reader all results, criticizing them, and allowing the reader to make his or her own judgement as to which value of the size and development of the shadow economy is more plausible. I have made this offer before and I am making it again.

REFERENCES

- Bajada, C. and Schneider, F. (2005). The Shadow Economies of the Asia-Pacific. *Pacific Economic Review*, 10(3), 379-401.
- Breusch, T. (2005). Estimating the Underground Economy using MIMIC Models, Australian National University. Working Paper, Canberra. Retrieved from: <http://econ.unimelb.edu.au/SITE/workshops/Breusch.pdf>.
- Dell'Anno, R. and Schneider, F. (2003). The Shadow Economy of Italy and other OECD countries: What do we know? *Journal of Public Finance and Public Choice*, 21(2-3), 97-120.
- Enste, D.H. and Schneider, F. (2006). Welchen Umfang haben Schattenwirtschaft und Schwarzarbeit? Ein Versuch zur Lösung des Rätsels, *Wirtschaftsdienst*, 2006/3, 185-191.
- Feige, E.L. (2016). Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? *Journal of Tax Administration*, 1/2016.
- Feige, E.L. and Urban, I. (2008). Measuring Underground (Unobserved, Non-Observed, Unrecorded) Economies in Transition Countries: Can We Trust GDP? *Journal of Comparative Economics*, 36(2), 287-306.
- Feld, L.P. and Larsen, C. (2005). Black Activities in Germany in 2001 and 2004: A Comparison Based on Survey Data, Study Nr. 12. Copenhagen: The Rockwool Foundation Research Unit.
- Feld, L.P. and Larsen, C. (2012a). *Undeclared Work, Deterrence and Social Norms: The Case of Germany*. Berlin/Heidelberg: Springer-Verlag.
- Feld, L.P. and Larsen, C. (2012b). Self-Perceptions, Government Policies and Tax Compliance in Germany. *International Tax and Public Finance*, 19, 78-103.
- Kazemier, B. (2006). Monitoring the Underground Economy: A Survey of Methods and Estimates. In Schneider, F. and D. Enste (Eds.), *Jahrbuch Schattenwirtschaft 2006/07. Zum Spannungsfeld von Politik und Ökonomie*, (pp. 11-53). Berlin: LIT Verlag.
- OECD, Organization for Economic Co-operation and Development. (2002). *Measuring the Non-Observed Economy: A Handbook* (pp.1-249). Paris, France: OECD Publications Service. 1-249.
- Pedersen, S. (2003). The Shadow Economy in Germany, Great Britain and Scandinavia. Copenhagen: The Rockwool Foundation Research Unit.
- Schneider, F. (2015). Schattenwirtschaft und Schattenarbeitsmärkte: Die Entwicklung der letzten 20 Jahre. *Perspektiven der Wirtschaftspolitik (PWP)*, 16(1), 3-25.
- Schneider, F. and Enste, D. (2000). Shadow economies: Size, causes and consequences. *Journal of Economic Literature*, 38/1, 73-110.
- Schneider, F. and Williams, C.C. (2013). *The Shadow Economy*. London: Institute for Economic Affairs (IEA).
- Schneider, F., Buehn, A. and Montenegro, C.E. (2010). New Estimates for the Shadow Economies all over the World. *International Economic Journal*, 24/4, 443-461.

- Schneider, F., Krstić, G., Arsić, M. and Randelović, S. (2015). What Is the Extent of the Shadow Economy in Serbia? In G. Krstić, and F. Schneider (Eds.), *Formalizing the Shadow Economy in Serbia*, (pp. 47-76). Heidelberg: Springer-Verlag.
- Thomas, J.J. (1992). *Informal Economic Activity*. Ann Arbor: The University of Michigan Press.
- UN, United Nations Economic Commission for Europe (2008). *Non-Observed Economy in national Accounts: Survey of Country Practices* (pp.1-340).Geneva, Switzerland.
- Williams, C.C. and Schneider, F. (2016). *Measuring the Global Shadow Economy: The Prevalence of Informal Work and Labour*. Cheltenham (UK): Edward Elgar Publishing.

PROFESSOR SCHNEIDER'S SHADOW ECONOMY (SSE): WHAT DO WE REALLY KNOW? A REJOINDER¹

Edgar L. Feige²

Abstract

Professor Schneider's "Comment" on my "Reflections" paper fails to answer the key critiques leveled against his MIMIC estimates of the shadow economy. His "Comment" only serves to reinforce the contention that his documentation is inadequate, that his normalization procedures are arbitrary and conceptually flawed, and that his explanation of the "serious calibration error" that reversed the trend of SSE earlier estimates is untenable. A recent attempt to reproduce his findings also concludes "that it is not possible to replicate Schneider's MIMIC indexes based on the documentation from the paper." In short, the evidence challenging the veracity of his MIMIC estimates of the worldwide shadow economy is so strong as to question their place in the academic, policy and popular literature.

INTRODUCTION

Let me begin by commending Professor Schneider's continual efforts to focus attention on a subject dear to both of us. What may not be apparent to the general reader of this debate, unversed in the esoterica of national income accounting or MIMIC models, are the important areas of agreement that Professor Schneider and I share. We both champion the idea that unobserved economies are an important yet often neglected component of macroeconomic analysis having significant implications for governance in both developed and developing economies. We agree that non-compliance with generally accepted institutional rules has significant consequences for efficiency, equity and growth. Moreover, we both acknowledge that measuring the nature, extent and trend of these non-compliant behaviors is both notoriously difficult and essential if we are to understand the importance of the phenomenon and empirically assess its far ranging consequences. Among the consequences of non-compliance, we both have expressed concern about the erosion of the tax base, leading to the growth of government debt and/or a decline of public service provision. We share a concern about the potential distortion of our nations' information systems, the inequities caused by its distributional effects and the consequences of corruption and illegal activities engendering the erosion of the moral fabric of society. Finally, we have both acknowledged the strengths and weaknesses of various measurement approaches. What then are our areas of disagreement?

As much as we agree on the importance of studying this phenomenon, we approach the problem of defining it very differently, and our taxonomic differences affect our assessment of appropriate methods of measurement and of our evaluation of the state of our current empirical knowledge. I contend that there is not one "shadow economy" but many unobserved economies and set forth a taxonomic framework, which identifies their complex interrelationships. Different empirical methodologies are required to estimate the composition, magnitude and trend of these different unobserved economies. I believe that the institutions

¹ Citation: Journal of Tax Administration, 2016, Vol. 2, No.2.

² Professor of Economics Emeritus, University of Wisconsin-Madison. [elfeige@wisc.edu]. I wish to acknowledge the thoughtful comments of W. Lee Hansen and the invaluable econometric insights generously offered by Trevor Breusch, whose dogged commitment to seeking the truth is exemplary.

whose rules are being violated have the greatest incentive, resources, information and expertise to measure both the extent of non-compliance, and its most salient consequences. Tax revenue authorities are best suited to measure the unreported economy and its resulting tax gaps, and national and international statistical agencies are most adept at obtaining exhaustive measures of national income and output. Building public confidence in these complex measures requires extensive and timely reporting, greater transparency, greater methodological consistency between countries and over time, and the inclusion of error ranges to reflect associated uncertainty³. Professor Schneider's (2016) "Comment" in no way contradicts these conclusions.

However, in contrast to my taxonomic approach, Professor Schneider makes no distinctions between the various unobserved economies, preferring to employ the catch-all phrase "shadow economy" to describe them all. After various iterations, he has now settled on a "narrow" definition that identifies Schneider's Shadow Economy (SSE) solely as the "underground" economy component of the non-observed economy (NOE). Since the NOE is composed of the "underground", "informal" and "illegal" economies, it follows that the ratio of SSE/NOE <1. Regardless of whether one accepts the figures presented in my Table 2 (Feige, 2016, p.20) or Schneider's (2016, Section 2.5, Tables 2 and 3), it is clear that the values of SSE typically exceed the estimates of officially measured NOE obtained from named representatives of national statistical agencies, by several orders of magnitude. Although Professor Schneider expresses amazement⁴ that the experienced statistical agencies of the Netherlands, Norway, Sweden and Turkey report such "remarkably low values" for their NOE's, his astonishment does not make their findings any less accurate.

We have both written extensively about the merits and shortcomings of various approaches to measuring unobserved activity. My own writings have decried the proliferation of Type 1 errors in the empirical economics literature (Feige, 1975) and questioned the reliability and robustness of early IRS audits and national accounting discrepancy estimates (Feige, 1989a). I described shortcomings of the Tanzi currency demand approach (Feige, 1986) and abandoning my own transactions method, I examined the implications of relaxing the assumptions of the simple currency ratio method in order to make them more realistic (Feige, 1989a). Feige and Urban (2003, 2008) demonstrated that various versions of the electric consumption method yielded unreasonable estimates of unrecorded income for transition countries. I have attempted to improve the accuracy of the general currency ratio approach with independent measures of currency velocity (Feige, 1989b) and with improved estimates of the amounts and locations of US dollars circulating abroad (Feige, 1996, 1997, 2003, 2012a, 2012b). After more than three decades of commitment to this area of research, I reluctantly felt compelled to conclude:

Given the shortcomings of conventional macro model estimates of the underground economy and the lack of transparency and consistency of NOE estimates, it is high time that the profession acknowledges how little we really know about the underground economies and their causes and consequences. (Feige & Urban, 2008, p.287)

The "Reflections" paper (Feige, 2016) reviewed and evaluated the subsequent progress made and the remaining challenges confronting national and international statistical agencies in their efforts to improve measures of the non-observed economy (NOE). I also reviewed the

³ Feige (2016, p.28)

⁴ Schneider (2016, Section 2.5).

MIMIC/CD method employed by Professor Schneider to estimate SSE for 162 countries for the period 1999-2007. Citing the critiques of prominent economists, I concluded that “SSE estimates suffer from conceptual flaws, apparent manipulation of the results and insufficient documentation for replication, questioning their place in the academic, policy and popular literature” (Feige, 2016, p.1).

Disregarding the incisive critiques of his MIMIC/CD methodology, Professor Schneider continues to champion its results “as the most reasonable estimates of the size of the shadow economy” (Williams & Schneider, 2016, p.36). The critical question this Rejoinder must address is whether Professor Schneider’s (2016) “Comments” provide any informative answers to the key challenges concerning the veracity of his work.

Specifically, does his “Comment” adequately address the charge that his documentation is insufficient and often inaccurate? Does he refute the conclusion that the meaning of the latent variable in his MIMIC model is so obscure as to question its relationship to any unobserved economy? Does he discredit the results demonstrating that his shadow economy estimates are multiples larger than they are expected to be? Does his “Comment” adequately explain the nature of his “serious calibration error” that forced him to reverse the trend and changed the size of all his worldwide shadow economy results?⁵ Does he explain the implications of finding that his MIMIC index is negative and the consequences of his arbitrary decision to add a constant term to make the index positive?⁶ As will be elaborated below, his oft-unsupported assertions and his inadequate and inaccurate responses only serve to reinforce my conclusions concerning the lack of veracity of his results. I shall document these charges, focusing on the major areas of concern: documentation, normalization, calibration and replication.

THE DOCUMENTATION ISSUES

Professor Schneider’s findings are controversial for a variety of reasons. As the “Reflections” paper documents, there is considerable uncertainty as to whether the MIMIC methodology employed is even capable of measuring the “narrow” concept Professor Schneider now defines as his Shadow Economy. Professor Schneider acknowledges how “notoriously difficult” it is to measure “economic activity that is deliberately hidden”, yet he claims to have succeeded in doing so worldwide for 162 countries for the period 1999-2007 to a reported accuracy of one decimal place. In his “Comment”, he characterizes the foregoing statement as an “attack” on him (Schneider, 2016, Section 2.2) rather than the simple statement of fact it represents. He goes on to contend, without documentation: “*I always state that these point estimates have a margin of error of +/- 15%*”(emphasis added).⁷

A more accurate statement by Professor Schneider might have read:

In Schneider & Williams (2013), I first announced, on page 30, but cited no evidence or documentation, that “*Estimates of the size of the shadow economy by MIMIC methods are generally thought to have a margin of error of +/-15 percent.*” On page 50, I reiterated that unsupported claim stating: “*As noted, the MIMIC estimates have an error margin of +/-15.0 per cent of their estimated value.*”

⁵ Schneider Buehn and Montenegro (2010b, p.1)

⁶ Schneider Buehn and Montenegro (2010b, p.18, footnote 24)

⁷ Feige (2016, p.18, footnote 24) cites the first and only mention of the margin of error in Professor Schneider’s published work.

When I discovered these surprising assertions, which I had never before encountered in Professor Schneider's papers, I wrote to him as follows:

Since I cannot find any calculation of these error margins explicitly in your book, can you explain to me how you calculated these error margins? Have you published these error margins in any previous publication that I may be unaware of? (Email-November 21, 2013)

To date, Professor Schneider has not answered this inquiry. However, his "Comment" introduces yet another confusing element concerning the magnitudes of SSE results. When confronted with the fact that the ratio of SSE/NOE is expected to be less than one for all countries, yet turns out to be uniformly considerably greater than one, Professor Schneider discovers a rationale for suddenly reducing all SSE estimates by 20%.⁸ Citing an obscure survey pertaining exclusively to Germany (Feld & Larsen, 2012, p.61), Schneider suggests that "one can make the assumption that roughly 20% needs to be deducted from these macro shadow economy measures to allow for legally-bought material which is already counted in official GDP." (Schneider, 2016, Section 2.5). The cited survey for Germany suggests that the figure could be as high as 25% and that illegal activities constituting another 27-30% are similarly included. Why not deduct 50% or 55% from the shadow economy measures? In addition, if these adjustments to SSE are reasonable now, then why has he never before applied them to his other published SSE estimates? These are just two added examples of Schneider's disconcertingly cavalier approach to documentation and data handling.

However, to comprehend fully the importance of "documentation issues" for assessing the credibility of Professor Schneider's findings, it is crucial to have a broad understanding of how SSE results are produced. SSE results depend on a two-step procedure. First, a MIMIC model is estimated, yielding a time dependent index ($\tilde{\eta}_{it}$) of the latent variable for each country. The index essentially determines the temporal trend of SSE. A second step is required to produce an estimate of the size of SSE (as a percent of recorded GDP) for a particular country at any moment in time (η^*_{it}). The MIMIC index ($\tilde{\eta}_{it}$) must now be scaled ("calibrated") to an exogenous measure of SSE for that particular country at a particular moment in time. Let (η^*_{i2000}) denote an exogenous estimate of SSE for country *i* in the year *t*= 2000. Then:

$$1) \eta^*_{it} = \tilde{\eta}_{it} / \tilde{\eta}_{i2000} \times \eta^*_{i2000} \quad ^9$$

Professor Schneider claims to obtain these η^*_{i2000} exogenous values from currency demand models for each of the 162 countries.¹⁰ The importance of documenting the provenance of each these 162 exogenous values becomes apparent when one recognizes how substantially these exogenous values influence Schneider's results. Recall that the MIMIC index essentially determines the temporal trend of SSE (within country variation) while the between country variation in SSE is due to the 162 exogenous values presumably derived from independent currency demand studies. A simple analysis of variance (ANOVA) applied to SSE temporal cross-country results suggests that the MIMIC procedure accounts for less than 2 percent of the total variance of Schneider's results with more than 98 percent of the variance due to the

⁸ See Schneider (2016, Tables 2 and 3, Section 2.5).

⁹ This is the country equivalent of the calibration equation (7) in Schneider, Buehn and Montenegro (2010a, p.19; 2010b, p. 18) and equation (2) in 2010c, p.453.

¹⁰ For a number of developing countries, Schneider calibrates his index to "base values originating from the year 2005 because of data availability" (Schneider, Buehn & Montenegro, 2010b, p. 18, footnote 24).

exogenous currency demand values. It seems that his much-vaunted MIMIC procedure contributes virtually nothing to the overall variation in SSE published figures.¹¹ It is clear that one cannot assess the veracity of Schneider's findings without investigating the provenance and reliability of the critical exogenous values that account for most of the variation in SSE results.

My requests to Professor Schneider for necessary data documentation date back to the year 2002 when he first presented DYMIMIC estimates for transition and OECD countries.¹² Since my earlier efforts were unsuccessful, I again requested documentation for the sources of his exogenous currency demand calibration values in 2008 and again in 2011, this time to Professor Buehn who initially expressed a willingness to provide me with the information after I met him at a professional conference. My data requests (repeatedly copied to Professor Schneider) included the following language:

I would be happy to have you simply answer the following question in detail:
How is each of the 162 benchmark values for the 162 countries specifically derived? (Emphasis original)

Your paper gives the impression that these estimates come from standard Tanzi type currency demand models that are described in Appendix 1 (p.37) but this is never clearly spelled out or referenced. This issue is crucial since these benchmark values establish most of the variation across countries and many of the substantive results of your paper rely heavily on cross-country variations. (Email July, 5, 2011)

As acknowledged in my "Reflections" paper (Feige, 2016, p.18 footnote 26) a year later, (July 16, 2012) Professor Buehn provided me with the raw data for the specifications listed in their (2010a) paper. However, there was no information concerning the sources of the key benchmark (calibration) values. After repeated requests to Professor Schneider, on January 2, 2013, I received his "Preliminary Documentation of the Size of the Shadow Economy in 171 Countries"¹³ whose introduction stated:

In this preliminary documentation the calibration figures of the size of the shadow economies in 171 countries when using the MIMIC approach and their exact source (literature reference) are shown, so that everyone can check, what figures have been used. For all these "starting" figures the exact sources are given in the literature review, so that everybody can trace them further back.

Unfortunately, Professor Schneider's 165-page document did not contain a single source identifying the original specific currency demand model from which his "starting" values were derived. More than 50 percent of the referenced studies were self-referential; whose source information provided nothing more enlightening than his standard notation, "Own

¹¹ This result is itself surprising and may be related to the mysterious manner in which "the MIMIC index has been adjusted to the positive range by adding a positive constant" (Schneider, Buehn & Montenegro, 2010b, p.18, footnote 24).

¹² The sources for his results were "Own calculations using DYMIMIC method" (Schneider 2002, Table 2, p. 7) and "Currency demand approach, own calculations" (Schneider, 2002, Table 3, p. 13). My unsuccessful efforts to obtain data and model specifications were documented in Feige and Urban (2008, p.288, footnote 1).

¹³ Professor Schneider's "Comment" (Schneider 2016, Section 3) incorrectly claims that on June 7, 2013 he sent me documentation for 179 countries. On that date, I did receive his revised documentation for 27 countries as indicated below.

calculations.” Most references were to papers written after Schneider had first published his calibration values and hence could not have been their source. Upon pointing out the uselessness of this document to Professor Schneider and requesting specific references identifying original sources for his year 2000 “calibration” values for just three or five countries, on June 7, 2013, I received a document entitled, “A Preliminary Documentation of the Size of the Shadow Economy in 27 Selected Countries” which explained:

This documentation has the purpose to provide the values of the shadow economy (in % of GDP) for 27 countries, which “served” as starting values of the calibration procedure for the MIMIC estimations of these countries, e.g. in the study by Schneider, Buehn and Montenegro (2010).

Once again, this “documentation” proved to be completely inadequate. In numerous cases, the referenced “sources” of the starting values referred to work published years after the values had already appeared in Schneider’s papers. For example, the “calibration” value for Cameroon initially appeared in Schneider and Klinglmair (2004, Table 7.1 p.41) yet its “source” was listed as Suslov and Ageeva (2009). In other instances, a currency demand model was not the source of the calibration value, nor did the cited “source” value bear any correspondence to the actual exogenous value Schneider used for calibration.¹⁴

On November 22, 2013, Professor Montenegro provided me with his raw data for 152 countries. My subsequent attempts to reproduce these data from original sources revealed numerous discrepancies as did my attempt to reconcile the Buehn and Montenegro datasets with one another. I tabulated these discrepancies in Excel files that I sent to Buehn, Schneider and Montenegro on November 26, 2013, with further questions concerning how the data were standardized. I never received any further clarifications concerning these data discrepancies or any replies to my requests concerning how the input data had been transformed. Professor Schneider’s “Comment” refers to the last email he sent me on November 25, 2013, and completely misquotes the text he sent me.¹⁵ I naturally ignored Professor Schneider’s disingenuous request for documentation from an unspecified paper instead of answering my inquiries to him.

The foregoing examples of inconsistencies, inaccuracies and ambiguities represent only a small sample of the numerous problems encountered in efforts to obtain information from Professor Schneider. They strengthen the conclusion that his lack of documentation concerning the provenance of the key “calibration” values, which explain virtually all of the variation of his results, casts a shadow on the veracity of his findings. Additional key gaps in the documentary record required for adequate replication pertain to data sources and transformations involving differencing, missing data, standardization procedures and choice of time periods. These gaps have precluded reproduction and replication of his findings for a decade.

¹⁴ For example, the “source” given for Serbia was Christie and Holzner (2004) whose reported estimate for the year 2001 was 19 percent. A “Household Income Tax Method” derived this figure. Schneider’s starting calibration value for Serbia was not 19 percent but incongruously 36.4 percent as published in Schneider (2007, Table 3.2.4, p. 19). Although Schneider’s “documentation” specifically referred to the Schneider, Buehn and Montenegro (2010c) study, that paper does not contain any shadow economy estimate for Serbia.

¹⁵ Schneider (2016, Section 3) states: “In a mail from November 25, 2013, I asked him ...”Would you be willing to send me similar detailed documentation of the econometric estimates of your last paper, **I mean the paper Feige & Urban published in 2008**, in order to estimate the size and development of the shadow economy?” The bolded words did not appear in his email to me. In their place were the words “**using your approach**”. I had no idea to which paper or which approach he was referring.

THE NORMALIZATION ISSUE

My "Reflections" paper called attention to the consequences of Professor Schneider's choice of normalizing indicators and normalizing coefficients. Professor Schneider acknowledges that his arbitrary choice of the sign of the normalizing coefficient determines the sign of the structural parameters of the causal variables, (Dell'Anno & Schneider, 2006, p. 5) and that he chooses a (-1) normalizing coefficient on GDP in order to obtain his desired result, namely, that tax rates and SSE are positively related. I have already argued that economic theory suggests the relationship to be either ambiguous or negative (Feige, 2016, p. 22). Having guaranteed his favored result with the arbitrary choice of the (-1) coefficient on GDP or average working time, Schneider goes on to misleadingly conclude with respect to his direct and indirect tax variables, "that both causal variables are highly statistically significant and have the expected positive sign in all equations." (Williams & Schneider, 2016, p.81). His results have the expected sign because he forced them to have the expected sign, not because his data supported his hypothesis.

In his "Comment", Professor Schneider also justifies his assumption that the normalizing coefficient on GDP should be (-1) because "an increase in shadow activity has a negative effect on official GDP development" (Schneider 2016, Section 4). Yet in Schneider (2009, p.1106), he concludes that the shadow economy and official GDP are complementary, requiring that his normalization coefficient should be chosen as (+1).¹⁶ Moreover, SSE and GDP must be positively related to one another to the extent that national income accountants capture the underground economy in recorded GDP.

THE CALIBRATION ERROR

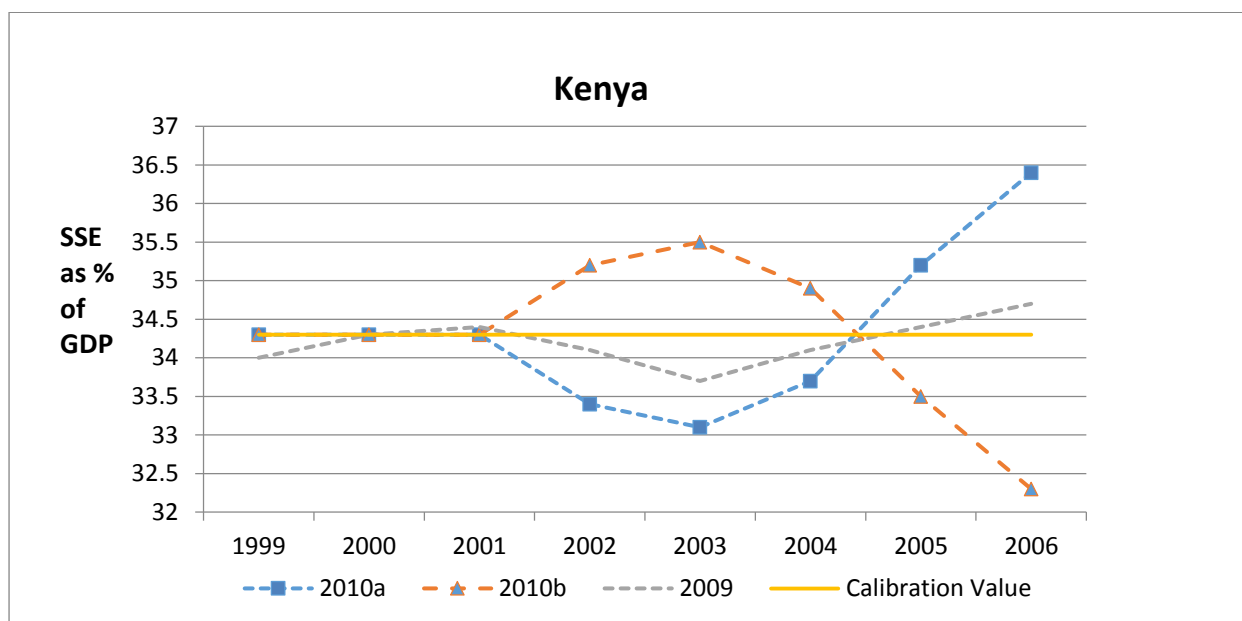
My "Reflections" paper focused attention on the "serious calibration error" (Schneider, Buehn & Montenegro, 2010b, p.1) that forced the authors to report completely different estimates of SSE for each of the 162 countries, reversing the trend of the shadow economy from their initial findings presented in Schneider, Buehn and Montenegro (2010a). Whereas their initial (2010a) paper reported a worldwide increase in SSE, the subsequent versions of the paper revealed a mirror image reversal in the worldwide trend. To be precise, let $SSE_a(s,i,t)$ represent Professor Schneider's shadow economy result published in the 2010a paper for specification s , country i in year t and let $SSE_b(s,i,t)$ represent the new SSE result published in 2010b. The following relationship holds for all s , i , and t :

- 1) $\{SSE_a(s,i,t) + SSE_b(s,i,t)\}/2 =$ Schneider's exogenously determined estimate of SSE for country i in year $t=2000$.

Thus, regardless of which specification they employ, the two curves representing the temporal path of SSE for every country are mirror images reflected about a horizontal line whose height is exactly the exogenous (calibration) estimate for the year 2000. Figure 1 displays the typical results using the example of Kenya, with the blue (square) line representing the initial faulty temporal path of SSE displayed in Schneider, Buehn and Montenegro (2010a), and the red (triangle) line representing the temporal path in Schneider, Buehn and Montenegro (2010b) after the "calibration error" was corrected.

¹⁶ Schneider (2009, p.1106) states, "government may not have a great interest to reduce the shadow economy due to the fact that: income earned in the shadow economy increases the standard of living of at least 1/3 of the working population, and between 40 and 50 % of the shadow economic activities have a complementary character, which means that additional value added is created, and this increases the official GDP"

Figure 1
Schneider's Shadow Economy Estimates for Kenya



The (green) dashed line displays the SSE estimates reported in Schneider and Buehn (2009). Since the 2009 results display a temporal path similar to the results of the faulty Schneider, Buehn and Montenegro (2010a) study, could the 2009 study also be subject to a similar “calibration error”?¹⁷

The author’s explanation of their mistake appeared in the opening footnote of the 2010b version that read:

Unfortunately the estimates of the original version (WPS 5356) needed to be revised due to a serious calibration error (sign switch). We apologize for this, especially as we now have in this version a negative trend for the size and development of the shadow economies over 1999 - 2007, which we did not have in the original version. (Schneider, Buehn & Montenegro, 2010b, p.1)

Their paper contained no further explanation concerning the nature of the calibration error. Professor Schneider’s “Comment” now informs us that:

Unintentionally, a sign error occurred in an Excel file. Unemployment was shrinking for almost all countries during the years 1999 to 2007 and, due to the mistake, the positive coefficient of unemployment was multiplied by -1. Hence, we found an increase in the shadow economy. Unfortunately, this mistake, which we did not notice immediately, occurred. Realizing the mistake, we immediately corrected it, updated the dataset and published a revised version of the (working) paper. (Schneider, 2016, Section 4)

¹⁷ These (2009) results were published in the Economics –Open Access-Open Assessment E-Journal that lists the Schneider and Buehn (2009) paper as its “most downloaded” article, having accumulated 26,396 downloads.

Empirical mistakes are readily made, and when speedily acknowledged, explained and corrected, cause little harm. However, the sign change on the unemployment coefficient that Professor Schneider now cites as their “mistake” could not have been responsible for the remarkable reversal of all of SSE results. Recall that the text in the original Schneider, Buehn and Montenegro (2010a), and the “corrected” Schneider, Buehn and Montenegro (2010b, 2010c) papers identically read:

the MIMIC model index of the shadow economies is calculated using the structural equation (1), i.e. by multiplying the coefficients of the significant causal variables with the respective time series. For the numerical example of specification 1 the structural equation is given as:

$$\tilde{\eta}_t = 0.14 x_{1t} - 0.06 x_{2t} - 0.05 x_{3t} - 0.27 x_{4t} \quad (6)$$

where, x_{1t} equals the size of government, x_{2t} and x_{3t} denote the business and fiscal freedom index, and x_{4t} represents GDP per capita.” [Schneider, Buehn & Montenegro (2010a, p18-19); (2010 b, p.17); (2010c, p. 453)]¹⁸

Note that the unemployment rate was omitted from their equation (6) since its coefficient was zero and statistically insignificant in the model for 98 developing nations.¹⁹ Since the unemployment variable was not included in the calculation of the MIMIC index, its supposed sign change could not have affected the size or trend of SSE results. Yet as illustrated by Figure 1, the authors systematically reversed SSE results for Kenya (and for all the other 97 countries) between the two versions of the paper.²⁰ This reversal of results occurred despite the fact that the coefficient estimates of the structural equation (6) remained the same and that the unemployment variable (the presumed source of the mistake) never entered into the calculation of the MIMIC index in either version of the paper.

How then could an error referring to the sign of a variable that was not involved in calculating the MIMIC index affect the size and trend of SSE results for all the countries in precisely the same symmetric fashion? Clearly, Professor Schneider’s explanation is completely implausible. However, if the “sign switch” on the unemployment variable is not the cause of the “serious calibration error”, then what was the nature of their “mistake”, how was it discovered and how was it remedied? Could a clue to the dramatic reversal of the SSE results, be contained in the mysterious addition to the footnote in the revised version, which stated that “the MIMIC index has been adjusted to the positive range by adding a positive constant” (Schneider, Buehn & Montenegro 2010b, p.18, footnote 24; 2010c, p.453, footnote 8).

What did Professor Schneider’s results look like before he arbitrarily added this mysterious constant to make his MIMIC index positive? Why did he find it necessary to add the constant?

¹⁸ Between versions, the authors have inexplicably renumbered all the specifications in the Tables. In the original (Schneider, Buehn & Montenegro 2010a, p.17), Specification 1 presents the results for “98 Developing Countries” whereas in the “corrected” (SBM 2010b, p. 16) version, Specification 1 refers to the results for “88 developed countries.” The “corrected” version (2010b) is in error since their equation (6) shows the parameters obtained for the “98 Developing Countries.” Schneider, Buehn and Montenegro (2010c) reverts to the same specification numbering as appears in the original version (SBM 2010a).

¹⁹ Schneider, Buehn and Montenegro (2010b, Table 3.1,p.16)

²⁰ Compare the results listed in Schneider, Buehn and Montenegro (2010a, Table 3.3.1, p. 20) with Schneider, Buehn and Montenegro (2010b, Table 3.3.2, p.21).

How large was the chosen constant and does its magnitude affect the size and trend of SSE results? Could the calibration error also have affected Schneider's earlier published shadow economy results, many of which were used in the derivative studies cited in Feige (2016, p.25-26)? These questions require answers if we are to understand the nature and full implications of the still unexplained "serious calibration error."

THE REPLICATION ISSUE

Reproduction, replication and robustness testing in economics should be quite straightforward when data sources, raw data, data transformations and statistical procedures are all fully documented and readily available. Although Professor Schneider presented MIMIC/CD results to the profession in 2002, to date, despite numerous attempts, only Trevor Breusch succeeded to replicate two of the early studies. His successful replications were the result of persistent detective work undertaken without assistance from Professor Schneider. Upon discovering the actual transformations and "benchmarking" procedures undertaken to produce the results, Breusch dismissed Schneider's complex applications of the MIMIC method, finding SSE results untenable.

My own efforts to reproduce the raw data from listed data sources were unsuccessful, as were my requests to obtain necessary clarifications from the authors. The most recent replication effort, of which I am aware, is a careful study undertaken by Marie-Astrid Maenhout (2016) who attempted to reproduce the raw data and replicate the derivation of the MIMIC index for specification (6) (25 High Income OECD countries) of the Schneider, Buehn & Montenegro (2010b) study.²¹ By limiting her focus to a single specification comprising the countries with the most readily accessible data and to the replication of the MIMIC index rather than the derivation of the exogenous "calibration" figures, Maenhout increased her chances for a successful replication. Because she was unable to reproduce the raw data for three of the eight variables employed in Specification (6), she proceeded with the replication effort using the raw data that Professor Buehn had supplied to me. Following the procedures outlined in the 2010b paper, she found that her parameter estimates of the causal model had the same signs as the published estimates except for the tax burden, which was significantly negative, suggesting that higher tax burdens were associated with smaller shadow economies. The most important variable in the Schneider, Buehn and Montenegro (2010b) paper, business freedom was the least important in her attempted replication. The least important variable in the (2010b) paper, the total tax variable had the second largest impact in the replication, albeit with the opposite sign.

In order to arrive at estimates of the size of SSE, Maenhout adopted Professor Schneider's exogenous calibration values of the shadow economy for the year 2000, and benchmarked the MIMIC index she had derived as described in Schneider, Buehn and Montenegro (2010b,

²¹ Ms. Maenhout's supervisor Professor Adriaenssens, on reading my "Reflections" paper, contacted me and requested the raw data that Buehn had sent to me. Adriaenssens' email read, "As we are only replicating the MIMIC index, and not the benchmarking procedure, we would benefit a lot from the raw data. We asked Prof. Schneider repeatedly, but all he managed to provide us are the final estimates. That is why we ask your help: could you provide us with these raw data? (Email –February16, 2016) quoted with the permission of Professor Adriaenssens (University of Leuven). On February 26, 2015 I sent Professor Adriaenssens the requested raw data as well as the Excel discrepancy worksheets I had prepared in attempting to reproduce and reconcile the Buehn and Montenegro datasets. After further requests, Schneider finally forwarded Buehn's data to Leuven. Ms. Maenhout's thesis summary concluded, "From the data reconstruction exercise I learned that it is not possible to reproduce the dataset based on the documentation from the paper."

equation 7). Whereas the Schneider, Buehn and Montenegro (2010b) estimates of SSE typically declined by less than two percentage points between 1999 and 2006, Maenhout's estimates fluctuated wildly with seven countries showing negative shadow economies for the year 2006. The 2006 estimates of SSE for Australia and Canada were -242% and -257% respectively. Maenhout (2016) concluded, "it is not possible to replicate Schneider's MIMIC indexes based on the documentation from the paper."

CONCLUSIONS

As much as we concur on the importance of studying unobserved economies, Professor Schneider and I approach the subject matter differently with regard to taxonomic and measurement issues. I prefer the analytic clarity of a taxonomic rule that results in definitional distinctions that correspond to empirically observable categories, whereas Professor Schneider employs the term "shadow economy" as a "catch-all". Professor Schneider interprets my "Reflections" paper as a personal "attack" on him, rather than the critical evaluation of both national accounting and macro methodologies it was intended to be. While I am personally sorry if my critiques of his MIMIC/CD method hurt his feelings, scientific accountability and informed public policy demand standards of documentation and replicability that transcend the feelings of any one individual.

Reproduction and replication are the gold standards of scientific inquiry. Replication serves to root out false claims and enables the profession to distinguish between "constructs", that is, results influenced by arbitrary decisions which bend the conclusions toward a researcher's prior opinions, and "estimates", namely, data-determined inferential outcomes obtained by applying accepted statistical procedures to coherently specified models. Not surprisingly, replication efforts are unlikely to be successful if the provenance of data sources are obscure, and the exact procedures followed in an analysis are inadequately documented.

My "Reflections" paper cited trenchant critiques of Professor Schneider's MIMIC/CD methodology, pointing out conceptual errors, non-robust results, undocumented and questionable data transformations, and concerns that arbitrary choices could substantially influence outcomes. As this "Rejoinder" documents, Professor Schneider's "Comment" is of no help in deflecting these concerns because it fails to resolve critical issues concerning documentation, normalization, calibration and hence, replication of his widely disseminated results. The most recent careful effort to replicate those results concludes, "it is not possible to replicate Schneider's MIMIC indexes based on the documentation from his paper." (Maenhout, 2016)

Professor Schneider's final offer that he and I write a joint paper "pointing to the differences between the various methods" used to measure the shadow economy fails to address the critical aforementioned issues. Over the past three decades, Professor Schneider has written so many papers and chapters repetitively describing the advantages and disadvantages of the various methods that I have stopped counting them.²² I doubt that the profession will benefit from yet another one, and therefore I must decline his magnanimous invitation for collaboration.

²² See for example: Alm, Martinez-Vazquez and Schneider (2004, Appendix A); Bajada and Schneider (2005, p.381-390); Schneider (1986, p. 645-649); Schneider (2005, Appendix A); Schneider, 2007, Appendix 1); Schneider (2009, p. 1114-1116); Schneider (2015, p 8-13); Schneider and Buehn (2009, Appendix A); Schneider and Buehn (2016, p 9-24); Schneider and Enste (2000, p. 91-99); Schneider and Enste (2002, Chapter 3); Schneider and Enste (2013, Chapter 3); Schneider and Williams (2013, p.27-31); Williams and Schneider (2016, Chapter 2).

Moreover, according to CollEc, Professor Schneider already holds the title of the world's top ranked economist on the scale of co-authorship centrality.²³ With eighty-five co-authors, he surely will not miss me as the eighty-sixth.

It is high time to move beyond recitations of the strengths and weaknesses of different approaches. If we are to determine which studies are replicable, robust and reasonable and which are undeserving of professional acceptance, we must penetrate procedural complexity to expose the intricate details of how results are attained. Only then can we assess whether the findings are legitimate data driven "estimates" resulting from commonly accepted inferential procedures. All the evidence to date challenges the veracity of Professor Schneider's worldwide shadow economy results. The inadequacies of his "Comment" to address this evidence only serves to reinforce the conclusions of my "Reflections" paper namely, that his findings suffer from conceptual flaws, apparent manipulation of results and insufficient documentation for replication, questioning their place in the academic, policy and popular literature.

²³ <http://collec.repec.org/rank/betweenness/1.html>

REFERENCES

- Alm, J., Martinez-Vazquez, J. & Schneider, F. (2004). "Sizing" the Problem of Hard to Tax. *AYSPS Conference: The Hard-to-Tax: And International Perspective*, Georgia State University. 1-72.
- Bajada, C. & Schneider, F. (2005). The Shadow Economies of the Asia-Pacific. *Pacific Economic Review*, 10(3), 379-401.
- Christie, E., & Holzner, M. (2004). Household Tax Compliance and the Shadow Economy in Central and Southeastern Europe. *wiiw Balkan Observatory*, Working Papers 038.
- Dell'Anno, R. & Schneider, F. (2006). Estimating the Underground Economy using MIMIC Models: A Response to T. Breusch's Critique (Working Paper No.0607) Linz, Austria: Department of Economics, Johannes Kepler University of Linz.
- Feige, E.L. (1975). The Consequences of Journal Editorial Policies and a Suggestion for Revision". *Journal of Political Economy*, 83(6), 1291-95.
- Feige, E.L. (1986). A Re-examination of the "Underground Economy" in the United States: A Comment on Tanzi". *International Monetary Fund Staff Papers*, 33(4), 768-781.
- Feige, E. L. (1989a). *The Underground Economies: Tax Evasion and Information Distortion*. Cambridge: Cambridge University Press.
- Feige, E. L. (1989b). Currency Velocity and Cash Payments in the US Economy: The Currency Enigma, *Munich Personal RePEc Archive Paper 13807*. Retrieved from <http://mpira.ub.uni-muenchen.de/13807/>.
- Feige, E.L. (1986). A Re-examination of the "Underground Economy" in the United States: A Comment on Tanzi. *International Monetary Fund Staff Papers*, 33(4), 768-781.
- Feige, E.L. (1996). Overseas Holdings of U.S. Currency and the Underground Economy. In S. Pozo (Ed.), *Exploring the Underground Economy* (pp. 5-62). Kalamazoo, Michigan: W.E. Upjohn Institute for Employment.
- Feige, E.L., (1997). Revised Estimates of the Size of the U.S. Underground Economy: The Implications of US Currency Held Abroad. In O. Lippert & M. Walker (Eds.), *The Underground Economy: Global Evidence of its Size and Impact* (pp. 151-208). Vancouver, BC: The Fraser Institute.
- Feige, E.L. (2003). The Dynamics of Currency Substitution, Asset Substitution and de facto Dollarization and Euroization in Transition Countries. *Comparative Economic Studies*, 45(3), 358-383.
- Feige, E.L. (2012a). New Estimates of U.S. Currency Abroad, the Domestic Money Supply and the Unreported Economy. *Crime, Law and Social Change*, 57(3), 239-263.
- Feige, E.L., (2012b). The Myth of the "Cashless Society: How Much of America's Currency is Overseas?" In *Deutsche Bundesbank Eurosystem, International Cash Conference 2012: The Usage, Costs and Benefits of Cash: Theory and Evidence from Macro and Micro Data*, (pp.119-151). Frankfurt am Main: Deutsche Bundesbank.
- Feige, E.L., (2016). Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? *Journal of Tax Administration*, 2(1), 5-41.
- Feige, E. L., & Urban, I. (2003). Estimating the Size and Growth of Unrecorded Economic Activity in Transition Countries: A Re-evaluation of Electric Consumption Method

- Estimates and their Implications. *William Davidson Institute Working Paper Number 636*. Retrieved from <https://core.ac.uk/download/pdf/3102894.pdf>.
- Feige, E. L., & Urban, I. (2008). Measuring underground (unobserved, non-observed, unrecorded) economies in transition countries: Can we trust GDP? *Journal of Comparative Economics*, 36(2), 287-306.
- Feld, L.P. & Larsen, C. (2012). *Undeclared work, Deterrence and Social Norms: The case of Germany*. Berlin/Heidelberg: Springer-Verlag.
- Maenhout, M.A. (2016). The Underground Economy: Evaluation and Replication of Schneider's Estimates. *Dissertation Summary, University of Leuven (KU Leuven)*, Brussels.
- Schneider, F. (1986). Estimating the Size of the Danish Shadow Economy using the Currency Demand Approach: An Attempt. *Scandinavian Journal of Economics*, 88(4), 643-668.
- Schneider, F. (2002). The Size and Development of the Shadow Economies of 22 Transition and 21 OECD Countries. *IZA DP No. 514*.
- Schneider, F. (2005). Shadow economies around the world: What do we really know? *European Journal of Political Economy*, 21, 598-642.
- Schneider, F. (2007). Shadow economies and corruption all over the world: New estimates for 145 countries. *Economics eJournal*, 2007-9.
- Schneider, F. (2009). Size and development of the shadow economy in Germany, Austria and other OECD countries. *Revue Économique*, 60(5), 1079-1116.
- Schneider, F. (2015). Schattenwirtschaft und Schattenarbeitsmarkt: Die Entwicklungen der vergangenen 20 Jahre. *Perspektiven der Wirtschaftspolitik* 16(1), 3-25.
- Schneider, F. (2016). Comment on Feige's paper "Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? *Journal of Tax Administration*, 2(2).
- Schneider, F., & Enste, D. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38, 77-114.
- Schneider, F., & Enste, D. (2013). *The Shadow Economy: An International Survey* (second edition). New York: Cambridge University Press.
- Schneider, F., & Buehn, A. (2009). Shadow economies and corruption all over the world: Revised estimates for 120 Countries. *Economics eJournal*, Version 2, 27 October.
- Schneider, F., Buehn, A., & Montenegro, C. E. (2010a). *Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007*. (World Bank Policy Research Working Paper 5356). Washington DC: The World Bank.
- Schneider, F., Buehn, A., & Montenegro, C. E. (2010b). *Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007 (Revised Version)* (World Bank Policy Research Working Paper 5356). Washington DC. : The World Bank.
- Schneider, F., Buehn, A., & Montenegro, C. E. (2010c). New Estimates for the Shadow Economies all over the World. *International Economic Journal*, 24(4), 443-461.
- Schneider, F., & Buehn, A. (2016). Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions. (*IZA DP No.9820*).

Schneider, F., & Enste, D. (2000). Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature*, XXXVIII, 77-114.

Schneider, F. & Klinglmaier, R. (2004). Shadow Economies around the World: What Do We Know? (IZA DP No. 1043).

Schneider, F. and Williams, C. (2013). The Shadow Economy. (IEA, Institute of Economic Affairs). London, UK: The Institute of Economic Affairs.

Suslov, N. & Ageeva, S. (2009). Influence of Energy Prices on the Size of the Shadow Economy: A Cross Country Analysis. (Project no-R06-1251, Working Paper No. 09/05E) Stockholm, Sweden: Economics Education and Research Consortium (EERC).

Williams, C. & Schneider, F. (2016). *Measuring the Global Shadow Economy: The Prevalence of Informal Work and Labour*. Cheltenham (UK): Edward Elgar Publishing.

REFLECTIONS ON THE MEANING AND MEASUREMENT OF UNOBSERVED ECONOMIES: AN EDITORIAL COMMENT.

Nigar Hashimzade, Chris Heady

The nature and extent of unrecorded economic activities is an important economic and policy issue for any country, and their measurement is a challenge for academic research. The exchange between Edgar Feige¹ ² and Friedrich Schneider³ highlights the difficulties facing researchers in this field. These difficulties range from data collection at the micro and macro level, to the choice of econometric techniques, and to the interpretation of the results.

An estimate of a hidden quantity, be it the output of an informal sector or unreported taxable income, is necessarily uncertain, and we cannot expect the level of accuracy in the measurement of the underground economy to be similar to that for the estimate of the formal economy. What we should expect is a well-grounded choice of the methodology and transparency in the description of data analysis, which would allow independent replication and assessment of the validity of the results.

We find Schneider's use of the MIMIC model in the measurement of underground economy unconvincing from a statistical perspective. As pointed out by Trevor Breusch⁴, this model is not appropriate for the analysis of macroeconomic data, because the endogenous links across macroeconomic aggregates are inconsistent with the MIMIC assumptions on the correlation structure in the set of variables. However, Schneider has provided a considerable amount of information about the model and how it was applied to the data in his work.

In contrast, Feige's work refers to estimates by the officers of the national statistical agencies (many of which were made available to him in personal correspondence), but does not describe the methods they employed. Here, we were unable to assess the estimates independently, because we have not been provided with sufficient information about the methodology. It is not impossible that the appropriateness of the methodology used by some countries is also questionable.

In rounding up this exchange we would like to express hope that the difficulties will not deter future research in the measurement of underground economy. In this challenging task, as in any empirical analysis, we need both further careful data analysis and transparency about the methods.

Nigar Hashimzade and Chris Heady

¹ Feige, E.L. Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? JOTA Vol. 2 No. 1 (2016): Special Issue: The Shadow Economy.

² Feige, E.L. Professor Schneider's Shadow Economy (SSE): What do we really know? A Rejoinder. JOTA Vol. 2 No. 2 (2016).

³ Schneider, F. Comment on Feige's paper 'Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"?' JOTA Vol. 2 No. 2 (2016).

⁴ Breusch, T. "Estimating the Underground Economy using MIMIC Models." JOTA Vol. 2 No. 1 (2016): Special Issue: The Shadow Economy.

THE INTERNATIONAL CONFERENCE ON TAX ADMINISTRATION

Simon James¹

The 12th International Conference on Tax Administration was organised by the School of Taxation and Business Law of the University of New South Wales (UNSW) Business School, and took place in Sydney on 31 March and 1 April 2016. The first of these conferences was held at the University of Newcastle NSW in 1994. In 1996, the second conference was held at the UNSW and subsequent conferences have taken place there every two years. This paper looks back at the original 1994 conference and considers how its original aims and contribution were reflected in subsequent conferences, and whether the subject matter's focus may have changed over the following 22 years, paying particular attention to the 2016 conference.

THE FIRST TAX ADMINISTRATION CONFERENCE

The aim of the first tax administration conference was to bring together the contributions of academics, practitioners and tax officials – an important feature which has continued ever since. It was entitled ‘Current Issues in Tax Administration’ and, although the word ‘international’ was not included in the early conferences, the first conference was already showing a strong international dimension, by drawing on experiences from a range of countries including not only Australia, of course, but also, in particular, India, Japan, New Zealand and the United Kingdom.

The conference was organised by Dr Ian Wallschutzky and took place at the University of Newcastle NSW on 7th and 8th April 1994. The 89 participants included 28 tax officials, (27 from the Australian Tax Office (ATO) and one from Revenue Canada) as well as 19 practitioners. The conference improved communications between the three groups, particularly between academics and tax officials. A total of 24 papers, which are listed in Appendix A, were presented in a single stream over the two days, and included one paper by members of the ATO and three by practitioners. Three members of the ATO presented a paper on ‘Managing the risks to the Revenue’, which was based on a preliminary analysis of the ATO’s Business Audit Programme, and sought to identify and weight major factors which, in practice, determined how that programme’s resources had been allocated between different taxpayer populations. The papers by practitioners consisted of: one on standards for the tax profession; another examining taxpayers’ rights and obligations; and a third which concerned penalties. The topic that attracted the most interest was tax compliance, which was the focus of seven papers. Almost all of the conference papers were concerned with issues that remain current, such as avoidance, compliance, legal processes and simplification.

THE MOVE TO UNSW AND SUBSEQUENT CONFERENCES

The 1996 conference moved to the Australian Taxation Studies Program (‘Atax’) at UNSW, where it has remained ever since. The conference organisers then came from Atax and have usually been any two of Chris Evans, Michael Walpole, Binh Tran-Nam, Margaret McKerchar

¹ University of Exeter. The author is grateful to Chris Evans and his colleagues at UNSW, Adrian Sawyer and Ian Wallschutzky for helpful comments on an earlier draft.

and Rodney Fisher. The 1996 conference kept the original title of ‘Current Issues in Tax Administration’ and the aim of including tax practitioners and officials. For example, the 1996 conference featured two papers on the Australian Taxpayers’ Charter – one by Michael Carmody, Australian Commissioner of Taxation, on ‘The ATO Perspective’, and the other by David Williams on ‘The Taxpayer’s Charter: A view from the Tax Profession’.

Each conference has included a different range of papers, but the topics examined have always been important aspects of tax administration. For example, the 1996 conference had five papers on compliance, four on tax simplification, four on tax dispute resolution, three on taxpayer service, three on taxpayer rights and obligations, and two on relative disclosure obligations and privilege. While the number of papers presented on each of these topics has waxed and waned over the years, tax compliance has been the focus of more papers than any another single subject area. As well as covering the usual topics, the conferences have also featured sessions devoted to topics that do not always receive a great deal of attention, such as tax collection (2004), legislative drafting (2006), benchmarking (2008) and building leadership (2014). Topics which have emerged or grown in importance over the past two decades, such as electronic filing and other technological developments, have also been included.

Over the years, a number of changes have been made. The title of the conference was changed to ‘International Conference on Tax Administration’ from 1998. It had always attracted overseas participants and the numbers grew (for instance, at the 2012 conference, 46 were from overseas out of a total of 130). Indeed, two of the small select group who have attended every conference are from overseas: Adrian Sawyer, from the University of Canterbury NZ; and Veerinderjeet Singh, originally at the University of Malaya and currently the Chairman of Taxand Malaysia. Furthermore, Adrian has presented a paper at all but one of the conferences. The conference has also attracted tax officials from an increasing number of countries, including Dave Hartnett, then permanent secretary for tax at HM Revenue and Customs (HMRC) in the UK, and several senior officials of the Internal Revenue Service (IRS) in the USA. A further change was the introduction of two streams for papers in order to include more within the traditional two-day format.

As the conferences continued they have often, but not always, had an overall theme. For example, the 6th Conference in 2004 had the theme of ‘Challenges of Globalising Tax Systems’ and the theme for the 10th, in 2012, was ‘Risky Business’. In addition to encouraging discussion about the usual wide range of topics, a conference theme can provide a helpful focus on a particular aspect. For instance, the theme of the 2010 conference was ‘Building Bridges’, and focused on relations between tax administrations and taxpayers. Sir Anthony Mason, who regularly opens the conference, described the concern of revenue authorities to change their image from possibly one of ‘grim-visaged, lantern-jawed tax gatherers’ to ‘considerate, understanding people whose goal is to assist taxpayers’ (Datt, Tran-Nam, & Bain, 2010, p. 3), and a number of papers at that conference examined how compliance might be affected by such developments. Whatever the theme of a conference, compliance has remained the subject area attracting the most papers. For instance, at the conference in 2010, the whole of one of the two streams was explicitly entitled ‘Tax Compliance’ and the other ‘Tax Administration’.

Finally, selected papers from the conference were formally published and edited by Evans and Greenbaum (1998), Evans and Walpole (2001), Walpole and Fisher (2003), Fisher and Walpole (2005), McKerchar and Walpole (2006), Walpole and Evans (2008) and Datt et al (2010). More recently, selected conference papers have been published in the *eJournal of Tax Research*. In

addition to papers appearing in the *eJournal of Tax Research*, many others have been published in top international peer-reviewed journals.

12th INTERNATIONAL CONFERENCE ON TAX ADMINISTRATION 2016

The 12th International Conference took place at the Crowne Plaza, Coogee Beach, Sydney. There were 10 speakers at plenary sessions and 22 papers were presented at the streamed sessions. The speakers, authors and their affiliations are shown in Appendix B. Once again, there was a good balance of contributions by academics, tax officials and practitioners, and a strong international dimension, with contributors from Australia, Austria, Indonesia, Japan, New Zealand, Turkey, UK, and USA. The papers received have been loaded onto the UNSW Business School (2016) website.

The first plenary session, at the beginning of the conference, consisted of presentations by: the Commissioner of Taxation, Australia; the Commissioner and CEO of Inland Revenue New Zealand; and the Assistant Minister of Finance for Tax Oversight at the Directorate General of Taxes, Indonesia. The Australian Commissioner, Chris Jordan, gave an address titled 'Better services and a better experience for Australians', which described trends in service delivery, including the reduction of red tape, user-driven design of the ATO's services, and differentiated, tailored engagement with taxpayers. The New Zealand Commissioner Naomi Ferguson's speech was titled 'Everything must change' and described the need for revenue authority change, not only in technology but also in processes, policy and revenue culture, in order to become more 'customer-focussed'. Puspita Wulandari then described the work of the Directorate General of Taxes in Indonesia and the challenges it faces. This was followed by a second plenary session in which Ali Noroozi, Inspector General of Taxation Australia, and Nina Olson, National Taxpayer Advocate, Internal Revenue Service, USA, described the work of their departments. They were followed by Shinichi Nakabayashi, Director of Administration, Management and Cooperation at the Asian Development Bank Institute, who gave a presentation on developing tax administration.

In the next plenary session, at the start of the second day, Duncan Bentley reviewed taxpayer rights in Australia twenty years after the introduction of the Taxpayers' Charter. Jeremy Sherwood, former Head of the Office of Tax Simplification (OTS) in the UK, examined tax complexity and the work of the OTS, and Ian Taylor, Chair of the Australian Tax Practitioners Board, focussed on the regulation of Australian tax practitioners. In the final plenary session at the end of the conference, Mark Chapman of H&R Block was concerned with the role of local tax agents in maintaining tax compliance. The session finished with a discussion involving a panel drawn from five nationalities. It consisted of Duncan Bentley, Eva Eberhartinger, Nina Olson, Adrian Sawyer and the present author, and it was chaired by Neil Warren.

Like the 1994 and subsequent conferences, the 2016 conference included more papers on compliance than on any other subject. Indeed, tax compliance took up one of the two parallel streams of specialist papers, with three of the four sessions devoted to various aspects of compliance, and the fourth with compliance costs and simplification. The other stream consisted of two sessions on tax administration and service delivery, and one each on taxpayer rights and dispute resolution. Some of the papers in this stream also had links to issues of compliance.

The first three sessions in the compliance stream included eight papers: two on behavioural issues, three with a strong international dimension and three on other aspects of compliance. One of the behavioural papers was by Michael Duggan of Inland Revenue New Zealand. His paper described research into mental tax accounting in relation to voluntary compliance and business systems. The concept of mental accounting is drawn from behavioural economics and describes the tendency for individuals to organise their money mentally into separate accounts as part of their personal financial management. It has generated some useful insights into financial behaviour and this paper presented survey results that found mental accounting is significantly related to compliance, as well as to a range of business systems and attitudes. Another behavioural dimension, this time regarding taxpayer perceptions, was used by Arifin Rosid, Chris Evans and Binh Tran-Nam. They examined whether, and how, perceptions of corruption may influence the compliance behaviour of personal income taxpayers, using evidence from Indonesia.

On international issues, Eva Eberhartinger and Matthias Petutschnig investigated the views of tax experts from practice around the world on the OECD Action Plan on BEPS (Base Erosion and Profit Shifting), contrasting the views of practitioners from BRICS countries (Brazil, Russia, India, China and South Africa) and from developing countries to those of practitioners from OECD countries. Ann Kayis-Kumar examined the effectiveness of thin capitalisation rules, and Agung Darono and Danny Ardianto presented a comparative study of CAATs (computer-assisted audit tools and techniques) in Australia, Finland, Germany, Indonesia and the USA. Other aspects of compliance were considered by: Michelle Drumbl, who explored taxpayer noncompliance with respect to the earned income tax credit (EITC) administered by the IRS; Neil Warren on the compliance risk that might be associated with electronic filing of tax returns using evidence from Australian personal income tax deductions; and Neni Susilawati on the role of school teachers in promoting tax compliance.

The final session of the compliance stream consisted of a paper on the methodological challenges of measuring compliance costs by Valmai Copeland and two on tax simplification. Tamer Budak and Simon James explored the applicability of the OTS Complexity Index to comparative analysis of the complexity of income tax and value added tax/goods and services tax in Australia, New Zealand, Turkey and the UK. Tamer Budak, Simon James and Adrian Sawyer reported on the experience of tax simplification in Australia, Canada, China, Malaysia, New Zealand, Russia, South Africa, Thailand, Turkey, the UK and the USA, and also examined the issue of distinguishing between necessary and unnecessary tax complexity.

In the other stream, the first two sessions were devoted to service delivery. The first paper was by Lyndall Crompton, Assistant Commissioner at the ATO, and outlined the ATO's strategies regarding international aspects of taxation, such as the BEPS Action Plan. The next paper was by Jo'Anne Langham of the ATO and Neil Paulsen of the University of Queensland, and described a new model of administrative effectiveness which combines existing knowledge from services management, public governance, engineering and psychology. The third paper, by Valmai Copeland and Virginia Burns of Inland Revenue NZ, presented findings on the use of software to prepare tax returns and digital media to communicate with the tax authorities.

The second session on service delivery included two papers. One was by Milla Setyowati, Fika Chandra, and Lita Khodariah, who provided an analysis of tax administration in Indonesia. The other was by Simon James and Andrew Maples, who examined the relationship between principles and policy in tax administration by analysing the UK capital gains tax (CGT) regime and suggested lessons regarding a proposal for the introduction of CGT in NZ.

The following session was on taxpayer rights and included three papers. John Bevacqua was concerned with the effects of enhancing taxpayer rights on tax compliance. Mathew Leighton-Daly proposed a model policy for the regulation of tax crime in Australia, drawing on his research and experience in practice as a barrister-at-law. The third paper was by Kalmen Datt, who evaluated the ‘naming and shaming’ approach that activists, the media and politicians sometimes take to the manner in which large corporations structure either themselves or their transactions to limit their tax liability. It was concluded that tax paid should be based on legal liability and ‘not be an ex gratia payment or attempt to appease unjustified, and often uninformed and vociferous criticism’.

The fourth session turned to disputes resolution. Melinda Jone focused on initiatives aimed at preventing or resolving disputes early in the process. These may involve a tax official trained in mediation techniques trying to reach an agreement with the taxpayer. In recent years, the revenue authorities in both Australia and the UK have formally adopted such arrangements, and this paper used dispute system design (DSD) principles to examine the UK system and offer recommendations for Australia. In the next paper, Ranjana Gupta drew attention to the role of tax practitioners which, of course, combines assisting clients with their financial affairs with a legal obligation to comply with the legal system, as well as professional responsibilities. Her paper reported the results of a survey of clients’ expectations and perceptions relating to tax practitioners in New Zealand. In the final paper in this session, Binh Tran-Nam and Michael Walpole reported some preliminary findings of a project on tax disputes, compliance costs and access to justice. The aims of the project are: to investigate access to independent tax dispute resolution, and to ascertain whether taxpayers with greater resources come out ahead and whether alternative dispute resolution (ADR) is effective.

The conferences up to and including the 2016 meeting have continued to reflect the aims, approach and coverage of the original in 1994. Their continuing success has been reinforced by a number of features, including attracting a wide range of academics, senior tax officials and practitioners, as well as many with international backgrounds. This not only provides a valuable multi-perspective approach to issues of tax administration but also helps to develop further links and relations between the different groups. The area which has continually attracted the most papers is tax compliance and there have been frequent illustrations of the benefits to be gained by combining insights from academics, officials and practitioners. Overall, it is clear the International Conference on Tax Administration continues to make a major contribution to advancing the study of tax administration, as well as always being a pleasure to attend.

REFERENCES

- Datt, K., Tran-Nam, B. & Bain, K. (2010). *International Tax Administration: Building Bridges*. Sydney: CCH Australia Limited.
- Evans C. & Greenbaum A. (Eds.), (1998). *Tax Administration: Facing the Challenges of the Future*. Sydney: Prospect Publishing.
- Evans C. & Walpole, M. (Eds.), (2001). *Tax Administration in the 21st Century*. Sydney: Prospect Publishing.
- Fisher R. & Walpole, M. (Eds.), (2005). *Global Challenges in Tax Administration*. Birmingham: Fiscal Publications.
- McKerchar, M. & Walpole, M. (Eds.), (2006). *Further Global Challenges in Tax Administration*. Birmingham: Fiscal Publications.
- UNSW Business School (2016, August, 3). 2016 International Conference on Tax Administration Papers. Retrieved from <https://www.business.unsw.edu.au/about/schools/taxation-business-law/conferences/tax-admin-papers>
- Walpole M. & Evans C. (Eds.), (2008). *Tax Administration: Safe Harbours and New Horizons*. Birmingham: Fiscal Publications.
- Walpole, M. & Fisher, R. (Eds.), (2003) *Tax Administration: Current Issues and Future Developments*. Sydney: Timebase.

APPENDIX A - AUTHORS AND PAPERS AT THE 1994 CONFERENCE

Peter Bardsley, La Trobe University.

Tax compliance research: An economic perspective on the research agenda.

Vicki Beyer, Bond University.

Tax administration in Japan.

Michael Blissenden, Senior Tax Writer Butterworths and University of Western Sydney.

Challenging s 167 assessments: A review of recent case law.

Cynthia Coleman and Lynne Freeman, University of Sydney and University of New South Wales.

The development of strategic marketing options directed at improving compliance levels in small business.

Gordon Cooper, Middleton, Moore & Bevins.

Standards for the tax profession

Graeme Cooper, Sydney Law School.

Incentives and strategic choices facing taxpayers under the self assessment system.

Thomas Delany and Kaye Emmerton, University of Southern Queensland.

An analysis of the variability in taxpayer responses to certain financial information items contained in income tax returns.

Abe Greenbaum, ATAX, University of New South Wales.

'David Jones Finance' and the review and appeal process.

Simon James and Ian Wallschutzky, University of Exeter and University of Newcastle NSW.

Should Australia adopt a cumulative withholding tax system?

Cliff Mancer, Massey University.

Tax simplification – predicaments and breakthrough for New Zealand.

Margaret McKerchar, Orange Agricultural College, University of Sydney.

A study of small business taxpayers in rural NSW.

Les Nethercott, Monash University.

Tax administration: A new era of compliance.

Lynne Oats, Dale Pinto and Pauline Sadler, Curtin University.

Penalties for tax agents in the brave new world of self assessment.

Bryan Pape, Exchequer Chambers.

Taxpayers' rights and obligations.

Greg Pearson, Ernst & Young.

What is 'reasonable' in the context of the penalty regime?

Jeff Pope, Curtin University.

Questioning a 'sacred cow': The Australian Taxation Office's lodgment programme/deadline system.

Graeme Purchas, University of Canterbury.

An examination of the tax practitioner's role in compliance.

Mahesh C. Purohit, National Institute of Public Finance and Policy, New Delhi.

Improving sales-tax management: A case study of India.

Cedric Sandford, University of Bath.

International comparisons of administrative and compliance costs of taxation.

Adrian Sawyer, University of Canterbury.

Binding rulings: Should New Zealand follow Australia's lead for once?

Barbara Smith, Deakin University.

Following the yellow brick road.

John Wickerson, Chris Grey and Nyree Goss, Compliance and Industry Research Unit,
Taxpayer Audit Group, Australian Tax Office.

Managing the risks to the revenue: A resource allocation perspective of the ATO's Business Audit Program.

Rob Woellner, University of Western Sydney.

Attitudes of ATO Auditors.

Frank Zumbo, UNSW.

Tax assessment and judicial review: Do the courts have a supervisory role?

APPENDIX B - SPEAKERS, AUTHORS AND PAPERS AT THE 2016 INTERNATIONAL CONFERENCE ON TAX ADMINISTRATION

Plenary Session 1

Chris Jordan - Commissioner of Taxation, Australia.

Naomi Ferguson - Commissioner and CEO of Inland Revenue New Zealand.

Puspita Wulandari - Deputy Director, Directorate General of Tax, Indonesia.

Plenary Session 2

Ali Nozoori - Inspector General of Taxation, Australia.

Nina Olson - National Taxpayer Advocate, Internal Revenue Service, USA.

Shinichi Nakabayashi - Director of Administration, Management and Coordination, Asian Development Bank Institute, Japan.

Plenary Session 3

Duncan Bentley - Swinburne University of Technology, Australia.

Jeremy Sherwood - Former Head, Office of Tax Simplification, UK.

Ian Taylor - Chair, Tax Practitioners Board, Australia.

Plenary Session 4

Mark Chapman - Director of Tax Communications, H&R Block.

Panel Discussion.

Papers in Streamed Specialist Sessions

Listed alphabetically by first author.

John Bevacqua, La Trobe University.

Taxpayer compliance effects of enhancing taxpayer rights: A research agenda.

Tamer Budak, Inonu University and Simon James, University of Exeter.

The applicability of the OTS Complexity Index to comparative analysis between countries: Australia, New Zealand, Turkey and the UK.

Tamer Budak, Inonu University, Simon James, University of Exeter and Adrian Sawyer, University of Canterbury.

International experiences of tax simplification and distinguishing between necessary and unnecessary complexity.

Virginia Burns and Valmai Copeland, Inland Revenue New Zealand.

Moving to digital by design: Better for customers, better for tax administration.

Valmai Copeland, Inland Revenue New Zealand.

Measuring compliance costs – methodological challenges.

Lyndall Crompton, Australian Tax Office.

ATO's focus on international taxation.

Agung Darono, Ministry of Finance Indonesia and Danny Ardianto, Monash University.
The use of CAATs in tax audits: A comparative study with lessons from international practice.

Kalmen Datt, University of New South Wales.
To shame or not to shame? That is the question.

Michelle Lyon Drumbl, Washington and Lee University.
Beyond polemics: Poverty, taxes and noncompliance.

Michael Duggan, Inland Revenue New Zealand.
Thinking tax: Mental (tax) accounting and voluntary compliance.

Eva Eberhartinger and Matthias Petutschnig, Vienna University of Economics and Business.
The scepticism of BRICS practitioners on the BEPS-agenda.

Ranjana Gupta, Auckland University of Technology.
Moderating influence of tax practitioner's explaining behaviour on the relationship between clients' service satisfaction and relationship commitment.

Simon James, University of Exeter and Andrew Maples, University of Canterbury.
The relationship between principles and policy in tax administration: Lessons from the United Kingdom capital gains tax regime with particular reference to a proposal for a capital gains tax for New Zealand.

Melinda Jone, University of Canterbury.
What can the United Kingdom's tax dispute resolution system learn from Australia? An evaluation and recommendations from a dispute systems design perspective.

Ann Kayis-Kumar, University of New South Wales.
What's BEPS got to do with it? Exploring the effectiveness of thin capitalisation rules.

Jo'Anne Langham and Neil Paulsen, University of Queensland.
Invisible taxation: Fantasy or just good service design?

Mathew Leighton-Daly, University of New South Wales.
A model policy for the regulation of tax crime in Australia.

Arifin Rosid, Chris Evans and Binh Tran-Nam, University of New South Wales.
Do perceptions of corruption influence personal taxpayer reporting behaviour? Evidence from Indonesia.

Milla Sepliana Setyowati, Fika Chandra and Lita Khodariah,
University of Indonesia.
Organizational Transformation of Indonesian Tax Administrator Authority.

Neni Susilawati, University of Indonesia.
Building tax culture in Indonesia: A case study of the role of school teachers in promoting tax compliance.

Binh Tran-Nam and Michael Walpole, University of New South Wales
Tax disputes, compliance costs and access to tax justice.

Neil Warren, University of New South Wales.
e-filing and compliance risk: Evidence from Australian personal income tax deductions.

REVIEW OF RECENT LITERATURE

Nigar Hashimzade¹, Lynne Oats²

This section provides a brief review of selected peer-reviewed publications from the first half of 2016 that investigate aspects of taxpayer and tax authority relations. It does not purport to be comprehensive, but rather gives a flavour of the variety of research studies from around the world. They are presented under broad headings, and in no particular order within those headings. The papers summarised here come from a variety of disciplinary backgrounds including economics, psychology, law and public policy.

TAX AUTHORITIES

Blank, J. D. & Osofsky, L. (2016) “Simplexity.”

To state it in simple words, simplexity occurs when a simple form hides a complex content. To comply with The Plain Writing Act of 2010, the IRS has to explain the tax law to individual taxpayers in plain language. Simplification can reduce the cost of compliance for taxpayers and the cost of administration for the tax authority, therefore potentially increasing tax revenue. However, according to the authors, the attempts by the IRS to use plain language in explanations of the complex tax law do not always benefit taxpayers. In particular, a plain language explanation of a law may fail to explain it in full. One negative consequence of this might be unequal outcomes for taxpayers, when sophisticated taxpayers will bear the lowest cost of simplexity. The authors offer a number of approaches that aim to address the drawbacks of simplexity while, at the same time, preserving its benefits. “Red-flagging”, for example, would require the tax authority to highlight the simplifications explicitly and to describe briefly alternative reasonable interpretations of the law.

Blank, J. D. (2017 – Forthcoming) “The timing of tax transparency.”

The focus of the paper is on the interesting and often overlooked aspect of fairness in tax law administration, namely, the timing of the public disclosure measures. The author argues that the accountability of a tax authority, such as the IRS, can be improved without undermining its ability to enforce the tax law, by requiring public accessibility for the documents related to taxpayers' tax affairs reflecting *ex ante* tax administration (for example, advance tax rulings or agreements), while preserving privacy for the documents reflecting *ex post* tax actions (audits or settlements).

Boll, K. (2015) “Proactive public disclosure: A new regulatory strategy for creating tax compliance?”

The paper describes a regulation strategy, termed “proactive public disclosure”, recently implemented by Danish Customs and Tax Administration (SKAT) against business engaged in VAT fraud. The strategy seeks to engage consumers in the enforcement of VAT compliance;

¹ Professor of Economics, Durham University

² Professor of Taxation and Accounting, University of Exeter.

namely, specific consumers who were buying products or services from non-VAT registered businesses were actively informed about that - although, in principle, any consumer could access the VAT registration information for any business in the public domain on the SKAT website. Thus, the strategy targets a third party rather than tax law offenders. The analysis in the paper shows that, in the context of the principles of Danish public administration, the strategy of proactive public disclosure does not violate the Duty of Confidentiality, but is incompatible with the Good Public Governance. The reason for the latter is that the information is used to expose tax offenders and to threaten them in order to induce compliance, and such a behaviour would not be compatible with good practices for private debt collectors.

Lederman, L. (2016) “The IRS, politics, and income inequality.”

In this note, Lederman describes what she believes to be the problems of inadequate funding of the IRS by the U.S. Congress and, overall, lack of strong congressional support of the IRS, which have led to the deterioration of enforcement statistics and a reduction of the IRS service to taxpayers. Lederman argues that a decline in enforcement and taxpayer service ultimately benefits those high-income and wealthy taxpayers who are more likely to engage in tax avoidance and tax evasion activities more, thus leading to growing inequality in the United States. She suggests that politicians favouring progressive income tax should appeal to fight against inequality in attempts to achieve a “more balanced treatment” of the IRS by the Congress.

TAX COMPLIANCE

Bornman, M. and Stack, E. M. (Lilla) (2015) “Specific rewards for tax compliance: Responses of small business owners in Ekurhuleni, South Africa.”

The purpose of this study was to investigate the effect of rewards on tax compliance behaviour. While a reward may encourage the desired behaviour, it can also crowd out an intrinsic motivation. To explore this issue, and to establish whether a specific form of a reward plays a role, the authors conducted a survey among small businesses in a metropolitan area of Ekurhuleni in South Africa. The survey participants were offered two scenarios of rewards by the tax authority for timely submission of tax returns: a lottery ticket, with a chance of winning a substantial prize, and a certificate with public acknowledgement of being a tax compliant business. The participants were asked a number of questions about the extent of their support of the reward. The results suggest that businesses support the recognition of their compliance by the tax authority, with little difference between the two scenarios. Younger entrepreneurs and smaller businesses appeared to be more in favour of the rewards; there was little difference in responses by education, race, and gender of the respondents, and by the type of industry.

Bruhn, M. & Loepnick, J. (2016) “Small Business Tax Policy and Informality: Evidence from Georgia.”

This paper explores the introduction, in 2010, of a preferential regime for small businesses in Georgia, and the way in which it affected both firm creation and tax compliance. The authors measure bunching below the eligibility threshold and find evidence of underreporting. The analysis is based on administrative data provided by the Georgia Revenue Service and provides insights into taxation and informality.

Casal, S. & Mittone, L. (2016) “Social esteem versus social stigma: The role of anonymity in an income reporting game.”

In this paper, participants in an experiment were subjected to different conditions in relation to the role of anonymity. The paper contributes to research on social pressure and tax compliance by measuring “(i) the tax evaders’ willingness to pay a fee to avoid publicity of their evasion...and (ii) the taxpayers’ willingness to pay a fee to identify tax evaders.” The authors conclude that “avoiding stigmatisation and social blame is particularly important for occasional offenders.” Furthermore, there is a misalignment between the value attributed to individual anonymity and that of others that can be exploited by tax authorities in framing deterrence strategies.

Choo, C.Y., Fonseca, M.A. & Myles, G.D. (2016) “Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment.”

The authors report on a tax compliance experiment with three subject pools: students, employees and self-employed taxpayers. They find self-employed taxpayers most compliant, the students least compliant. The experiment bridges two streams of prior research: laboratory experiments and randomised controlled trials. The main finding is that there is a stark difference between students and non-students, which raises important questions about the efficacy of using students for tax compliance experiments.

Hashimzade, N., Myles, G. & Rablen, M.D. (2016) “Predictive analytics and the targeting of audits.”

This paper demonstrates the use of predictive analytics in an agent-based model of a social network that governs the interaction and transmission of information between taxpayers. Taxpayers make an occupational choice between being self-employed or employed, where the latter choice allows no possibility of non-compliance. The predictive analytics use data from tax returns and from the outcomes of past audits, and it is shown that they produce more revenue than a strategy of random audits. One conclusion of the paper is that it may be appropriate for tax authorities to condition audit strategies on both reported income and also occupation.

Manca, M. (2016) “The new Italian cooperative compliance regime.”

The article gives a comprehensive description of a new regime introduced in the Italian domestic tax system as of January 1, 2016. This regime is based on the new approach to the relationship between taxpayers and tax administrations - that of trust and cooperation. The background, as outlined in this article, consists of several projects and initiatives, such as, for example, the OECD framework on cooperative compliance and the OECD Action Plan on BEPS. The Legislative Decree No. 128 (5 August 2015) specifies duties for the Italian tax administration and taxpayers that would help improve transparency and impartiality in tax compliance. The regime will ensure faster domestic ruling procedures and the possibility of reduced penalties when a taxpayer and tax authority disagree in the interpretation of a tax issue. The Decree also refines the boundaries of the judicial doctrine of the abuse of rights which, according to Manca, was often criticised by tax practitioners.

Onu, D. & Oats, L. (2016) ""Paying tax is part of life": Social norms and social influence in tax communications."

It has long been recognized that an individual's tax compliance decisions are often influenced by their social environment. But what exactly do taxpayers discuss and how do their communications lead to changes in compliance? Onu and Oats explore these issues using the discourse analytic approach for an in-depth analysis of online discussions about income tax among software developers. Five broad types of communications have been identified in the sample of 120 conversations: giving and receiving information about tax rules and procedures; communicating social norms of compliance; communicating benefits of compliance; urging to comply by direct persuasion; and warning about penalties or threatening with reputation loss. These findings help to shed light on the process of formation of social norms in tax compliance.

Randlane, K. (2016) "Tax Compliance as a System: Mapping the Field."

In this paper, the author attempts to approach the compliance puzzle in a holistic manner. She provides a systematic analysis of the compliance literature, not only that emanating from economics and psychology but also more recent ethnographic perspectives, in particular Boll (2014); charting its changing focus from deterrence to non-economic considerations. Viewing tax compliance as a system, the author develops an abstract model and suggests we move away from a viewing tax compliance as only being associated with tax evasion towards a comprehensive systems analysis, in order to develop more robust administrative strategies.

Satterthwaite, E. A. (2016) "Can audits encourage tax evasion? An experimental assessment."

This experimental study has investigated dynamic effects of tax audits. How does a taxpayer respond to an audit – does live experience of an audit improve or deteriorate compliance in the future? The economic psychology literature mentions two competing predictions; the “target effect” and the “bomb-crater effect”. According to the former, post-audit compliance improves because an audited taxpayer will feel targeted by the tax authority and so will consider themselves to be more likely to be audited in the future. On the contrary, the latter implies that being audited again in the future is less likely, as a bomb is unlikely to hit again the same place, and so post-audit compliance deteriorates. Satterthwaite reports that in this study, based on the survey of 201 U.S. resident workers on Amazon’s Mechanical Turk, both effects have been observed, depending on how the audit was framed. Specifically, there was strong evidence of the bomb-crater effect (a drop of 8 percentage points in average compliance rates) after the first audit, which was described as “random” to the participants. However, informing the participants that evasion will trigger future audits according to some “endogenous audit rule” resulted in improved compliance (a 4 percentage point increase). At the same time, while the second audit led to a “dramatically” improved compliance rate, the effect of a string of consecutive audits on average compliance was inconclusive, suggesting that repeated audits might not be cost-effective.

TAX AVOIDANCE

Van der Pas, J. (2016) “Improving the Chinese Anti-Avoidance Rule: A Comparative and Functional Approach.”

The author observes the spread of GAARs and examines developments in China in light of developments in other countries. The paper takes a comparative functional approach that looks not only at the ‘law on the books’ but also ‘law in action’, bringing into the analysis cultural and institutional differences, as well as case law and scholarly opinions. The Chinese GAAR is evaluated by reference to the Netherlands and the UK, and the author finds that while the approaches in these three jurisdictions differ, there are, nonetheless, some common themes. In the Chinese setting, the balancing of taxpayer rights against the interests of tax authorities is ‘easily overlooked’, and the discretionary powers of the Chinese tax authorities lead to an ‘uneven playing field’.

TAX EVASION & CORRUPTION

Alm, J., Martinez-Vazquez, J. & McClellan, C. (2016) “Corruption and firm tax evasion.”

There is a large body of literature on each of the subjects of corruption and tax evasion, yet the relationship between these is not well understood. Corruption, in this paper, relates to the bribing of tax officials and, in particular, the authors consider how the potential for bribery affects firms’ tax evasion decisions. The authors use the World Enterprise Survey, as well as the Business Environment and Enterprise Performance Survey, and use an estimation approach, with propensity score matching as a robustness check. They find that corruption of tax officials is a significant determinant of tax evasion and further that larger bribes result in higher levels of evasion, which underscores the importance of honesty in tax administration.

Litina, A. & Palivos, T. (2016) “Corruption, tax evasion and social values.”

In this paper, the authors propose a theoretical explanation for the relationship between political corruption and tax evasion. They note that the main conclusions of prior literature are that non-evading individuals view tax evasion as immoral and that those with friends who evade tend to do so themselves. The literature also shows that compliance is higher where there is a stronger feeling of social cohesion and that social norms are important in evasion decisions. Amongst other things, the authors conclude that the presumption that trust in tax authorities impacts positively on tax compliance may not hold in societies with a high level of corruption, making it difficult for honest leaders trying to deal with both corruption and tax evasion.

TAX PRACTITIONERS

Finley, A. R. & Stekelberg, J. (2016) “The economic consequences of tax service provider sanctions: Evidence from KPMG’s deferred prosecution agreement.”

The authors investigate whether KPMG’s business of selling tax services suffered after it acknowledged, in June 2005, an “unlawful conduct” in marketing tax shelters to individual clients, and entered a deferred prosecution agreement (DPA) with the Department of Justice. Under the DPA, KPMG was subjected to higher scrutiny during the first three years and agreed to comply permanently with stricter requirements on tax avoidance strategies compared to

other tax service providers. The authors find that after the DPA, the clients were 1.4 times more likely to stop or reduce purchases of tax services from KPMG than from other Big 4 accounting firms. However, the remaining KPMG clients showed no change in tax avoidance behaviour compared to other Big 4 clients. Thus, the sanctioned tax service provider suffered negative economic consequences, whereas there was little or no observable tax cost to the clients who continued to buy tax services from a sanctioned provider.

Frecknall-Hughes, J., Moizer, P., Doyle, E. & Summers, B. (2016) “An examination of ethical influences on the work of tax practitioners.”

This paper addresses two interesting questions arising from the recent debate about the allegedly unethical tax avoidance practices of large multinational companies and the role of tax professionals. The first question concerns the conceptual frame within which tax professionals operate and its effect on the priority given to different ethical considerations. The second question concerns the effect of the tax context on ethical reasoning relative to a more general social context. The Defining Issues Test is used to investigate the use of a consequentialist and deontological approach to moral dilemmas by tax practitioners.

TAX LITIGATION

Freedman, J. (2016) “UK institutions for tax governance: Reviewing tax settlements.”

With the reference to an earlier note by the same author, this article continues a review of the UK tax institutions in the context of sensitive or large settlements. Specifically, the author outlines the issues related to the introduction in 2012 of the post of the Tax Assurance Commissioner and the measures taken towards further enhancement of transparency in the HMRC operations.

INTERNATIONAL CO-OPERATION

Hakelberg, L. (2016) “Coercion in international tax cooperation: Identifying the prerequisites for sanction threats by a great power.”

In October 2014, 51 jurisdictions, including most offshore tax havens, signed a multilateral competent authority agreement (MCCA) under which they resolved to apply, from September 2017, the common reporting standard (CRS) developed by OECD and to set up a peer-review process of monitoring the implementation of the CRS. This article offers an explanation for this major step towards the eradication of tax evasion and tax avoidance. The author argues that this agreement is, in essence, a redistributive cooperation in which there are winners and losers among the participant governments. As such, it was induced by a credible threat of sanctions by a “great power” (defined as a jurisdiction with large internal markets, and less dependent on foreign investment and trade than smaller countries). Hakelberg’s analysis of tax negotiations at OECD level in 1996 - 2014 appears to support two hypotheses about the enforcement of international tax cooperation by sanction threats: (H1) “*domestic constraints preventing a shift of the tax burden from capital to labour or consumption*”, and (H2) “*scope for redistributing cooperation benefitting great power banks or multinationals*”.

REFERENCES

- Alm, J., Martinez-Vazquez, J., & McClellan, C., (2016). "Corruption and firm tax evasion." *Journal of Economic Behavior and Organization*, 124, 146-163.
- Blank, J.D., & Osofsky, L. (2016). "Simplexity". *Emory Law Journal*, 66. (forthcoming)
- Blank, J.D. (2017). "The timing of tax transparency". *Southern California Law Review*, 90 (forthcoming)
- Bornman, M., & Stack, E. M. (Lilla) (2015). "Specific rewards for tax compliance: Responses from small business owners in Ekurhuleni, South Africa." *eJournal of Tax Research*, 13(3), 799-818.
- Bruhn, M., & Loeprick, J. (2016). "Small Business Tax Policy and Informality: Evidence from Georgia." *International Tax and Public Finance*, 23(5) 834-853. doi: 10.1007/s10797-015-9385-9
- Casal, S., & Mittone, L. (2016). "Social esteem versus social stigma: The role of anonymity in an income reporting game." *Journal of Economic Behavior and Organization*, 124, 55-66.
- Choo, C.Y., Fonseca, M.A., & Myles, G.D. (2016). "Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment". *Journal of Economic Behavior and Organization*, 124, 102-114.
- Hakelberg, L. (2016). "Coercion in international tax cooperation: Identifying the prerequisites for sanction threats by a great power". *Review of International Political Economy*. doi: 10.1080/09692290.2015.1127269
- Finley, A.R., and Stekelberg, R. (2016). "The economic consequences of tax service provider sanctions: Evidence from KPMG's deferred prosecution agreement". *Journal of the American Taxation Association*, 38(1), 57-78.
- Freedman, J. (2016). "UK institutions for tax governance: Reviewing tax settlements". *British Tax Review*, Issue 1, 2016, 7-12.
- Frecknall-Hughes, J., Moizer, P., Doyle, E., and Summers, B. (2016). "An examination of ethical influences on the work of tax practitioners". *Journal of Business Ethics*.
- Hashimzade, N., Myles, G., & Rablen, M.D. (2016). "Predictive analytics and the targeting of audits." *Journal of Economic Behavior and Organisation*, 124, 130-145.
- Lederman L. (2016). "The IRS, politics, and income inequality". *Tax Analysts: Tax Notes*, March 14, 1329-1333.
- Litina, A., & Palivos, T. (2016). "Corruption, tax evasion and social values." *Journal of Economic Behavior and Organisation*, 124, 164-177.
- Manca, M., (2016). "The new Italian cooperative compliance regime". *European Taxation*, April 2016, 152-160.
- Onu, D., and Oats, L. (2016). "'Paying tax is part of life': Social norms and social influence in tax communications". *Journal of Economic Behaviour & Organization*, 124, 29-42.
- Randlane, K. (2016). Tax Compliance as a System: Mapping the Field. *International Journal of Public Administration* 39(7), 515-525. doi:10.1080/01900692.2015.1028636
- Satterthwaite, E.A. (2016). "Can audits encourage tax evasion? An experimental assessment". *Florida Tax Review* (forthcoming)

Van der Pas, J. (2016). "Improving the Chinese Anti-Avoidance Rule: A Comparative and Functional Approach." *World Tax Journal*, 80-100.