CONCLUSIONS PAPER

How to Catch a Tax Cheat

THE POWER TO KNOW.

Contents

Featuring

Shaun Barry, Principal Solutions Architect, SAS Security Intelligence Practice
Jon Lemon, Solutions Specialist, SAS Federal

Introduction

Tax evasion is the largest economic crime in the world (in terms of monetary loss), costing trillions of dollars to governments around the globe. A 2011 study by The Tax Justice Network estimates that on a global scale, total tax evasion is in excess of US$3.1 trillion, or about 5.1% of world GDP.[1] And that's just the known tax evasion and noncompliance; it doesn't include the underground economy and cash businesses.

The per-capita figure is similar in Canada, representing about $45 billion a year. The European Union loses an estimated €200 billion in value-added tax alone, not including corporate or excise taxes. Tax evasion represents 4 to 5 percent of gross domestic product in Sweden and Japan.

"In some countries, such as Brazil or Pakistan, tax evasion is a virtual sport," said Shaun Barry, Principal Solutions Architect in the SAS Security Intelligence Practice. "Pakistan estimates that it collects only $.30 of every dollar it should collect in government revenues. For Brazil, it's only about $.60 of every tax dollar."

What would it mean if these governments could identify and collect all the tax revenues they're supposed to? "Hypothetically, if the IRS had been able to solve the known tax gap of $385 billion each year from 2000 to 2014, an $8.6 trillion deficit would be whittled down to $1.5 trillion," said Barry. "Granted, $1.5 trillion is still a lot of money, but if we had been able to solve the tax gap problem back then, public debate in Washington would be substantially different than it is today."

A webinar sponsored by SAS and hosted by the Association of Certified Fraud Examiners examined this topic. What tax evasion schemes are prevalent and how can they be detected and prevented? Recent advances in computing and analytical technology have given tax administrators new power – capabilities and lessons learned that apply to many different types of fraud across other industries as well.

> "Some of the lessons that tax administrators and government officials have learned over the last 10 years in this area are highly applicable to different types of fraud across industries."
>
> Shaun Barry,
> Principal Solutions Architect, SAS Security Intelligence Practice

Common Forms of Tax Evasion

The first step in resolving the tax gap is to identify the components of noncompliance:

- **Filing compliance.** If you're legally obligated to file a tax return, do you actually meet that obligation, or are you a nonfiler?
- **Reporting compliance.** If you do file your tax return, do you tell the truth about the figures on that tax return, or are you an underreporter?
- **Payment compliance.** If you file a truthful return that says you owe a certain amount in taxes, do you pay it?

"These components matter because the types of approaches you take to detecting noncompliance and evasion will vary widely because the underlying business problems are different," said Barry.

Where there is information reporting, governments have a powerful tool for stemming the tide of tax evasion. Third-party information can verify that a taxpayer is telling the truth – or not – on a tax return. Where there is substantial information reporting –

[1] *The Cost of Tax Abuse; A briefing paper on the cost of tax evasion worldwide*, The Tax Justice Network, November 2011.

such as salaries and withholding reported on employer W-2 forms – compliance is high and fraud is estimated to be as low as 1 percent. But in areas with less information reporting, fraud and noncompliance can soar to more than 50 percent. For example, the IRS estimates that about $150 billion a year in small business income is misreported on Schedule C forms.

Several types of schemes have been particularly difficult to catch due to lack of information reporting, said Barry.

- **Identity theft.** Taxpayers submit false or fake returns using a stolen identity to get refunds or rebates, or steal identities of others to claim them as dependents to increase their own tax credits or deductions.
- **Creative "situs."** Business entities or individuals place economic activities in low – or no-tax jurisdictions – tax havens such as the Cayman Islands or (until recently) Ireland – in ways that are clearly designed to evade complying with tax laws.
- **Shell companies and networks.** A complex network of business entities – partnerships, limited liability corporations (LLCs) and other shell companies – is set up to move financial transactions and income among entities and into and out of various jurisdictions.
- **Underground economy.** Your friendly neighborhood loan shark, ticket scalper, drug dealer and other cash-intensive businesses operate in a gray area where little or no income is reported.

If governments cannot rely on information reporting to help detect these tax evasion opportunities, what can they do?

Combat Tax Fraud With Analytics

Jon Lemon, Solutions Specialist with the SAS Federal practice, described analytical techniques – some simple, some sophisticated – that can be applied to the data to help filter out some of those fraudsters:

- **Business rules** to identify known inconsistencies and disconnects.
- **Predictive modeling** to detect patterns associated with known fraud.
- **Anomaly detection** to find unusual attributes or transactions.
- **Database matching** to identify potential cases from nontax information.
- **Link analysis** to connect the dots among related entities.

Business Rules

"Business rules are a good first line of defense, commonly used to find less sophisticated schemes or mistakes in tax returns," said Lemon. SAS has a library of rules – some patented, some patent pending – based on years of experience working with federal, state, local and international tax agencies. For example, rules can catch cases where a person claims a dependent but doesn't offer a valid Social Security Number, or where deductions are out of line with a person's tax bracket. "But if someone actually is trying to steal from you and rules can catch it, then they're not very sophisticated in their thinking," said Lemon.

Predictive Modeling

Predictive models can predict the behavior of a taxpayer and identify suspicious patterns and trends. For example, suppose a tax administration is likely to approve a transaction/deduction under $10,000 for a given scenario. Anything above that threshold will be denied or receive extra scrutiny.

A fraudster could submit at different levels in an attempt to determine where that threshold is set. Perhaps the first submission will be for $3000, which is approved. Maybe the next time it is $7000, also approved. Then $9000, also approved. Finally the person will try $12,000, which is denied. Now the taxpayer knows where the threshold is and how high the figure can be while still avoiding unwanted scrutiny or having the exemption denied. In isolation, each of these transactions might look innocuous, but taken together, they show a clear stair-stepping pattern designed to test the system.

"When you're looking at a sea of data, millions and millions of tax returns over a span of many years, it's very difficult to see how those individual transactions relate to each other," said Lemon. "With predictive modeling that knows the patterns that have been established by previous evaders, we can identify suspicious activity and score that behavior higher."

Anomaly Detection

Is the information on a taxpayer's submission typical for his/her income level, family status or occupation – or more typical of fraudsters? Anomaly detection finds possible tax evasion in two different ways:

- Establish a **peer group** – perhaps by geographic area, income bracket or families of a certain size – and compare the subject to the peer group. "If that person is taking deductions that are maybe five, six or seven standard deviations outside of the norm, they're going to get a higher score than a person whose deductions are just one or two standard deviations off the norm," said Lemon.
- Build **profiles** of what normal and abnormal behavior look like. "If we see a sudden change in behavior, this could be indicative of identity theft," said Lemon. "For example, if the current return has very little in common with the filer's history, we'd want to examine whether there is a legitimate reason for that before issuing a large refund."

Database Searches

There's a lot of value in mining both internal and external data sources to cross-match and verify (or refute) information that is self reported by the tax filer. This could include internal sources such as known bad lists, extra-agency sources such as death and incarceration records, and third-party commercial sources that can provide address or identity verification. For example, a person who appears on a state list for unpaid or delinquent taxes might not be paying federal taxes either. A person showing evidence of financial difficulty on a credit report might be less likely to be wholly truthful in reporting income. The possibilities are endless for using data and text mining of relevant and appropriate data sources to find inconsistencies in a taxpayer's claims or signs of possible noncompliance.

Link Analysis

"Link analysis is one of the more powerful ways to find tax evaders," said Lemon. You can identify networks of people linked by address, business relationships, family relationships or maybe by bank account numbers. For example, you might find a case where a single tax preparer is linked with 100 different customers who have family members who are tax evaders. That information might direct your attention.

"The ability to detect the network and then put the network through analytical techniques makes the net even stronger," said Lemon. "One person alone might be able to fly under the radar, but as a network, it's much, much more difficult to go unnoticed. If we see one taxpayer flagged for fraud, then look to see who the tax preparer was, we may find five or six other taxpayers associated with that tax preparer, which gives us some clues where to look further."

> "Link analysis is an extremely powerful tool to not just find one or two tax evaders but to also get to the root of the problem and discover more sophisticated schemes at the network level."
>
> Jon Lemon,
> Solutions Specialist,
> SAS Federal

How Much Data Can Be Crunched in Less Than a Second?

Big data analytics and high-performance computing have redefined the possible. Processing tasks that were once inconceivable are now reality. Here's an example: "We're working with a large banking institution that scans every credit card transaction at the point of sale, every day," said Barry. When the customer swipes a credit card, data is sent to a SAS system, which scores the transaction using analytic techniques and returns one of three signals:

- Approve the transaction if the cardholder's identity is valid.
- Deny the transaction if the purchaser is not the valid cardholder.
- Approve the transaction but flag it for a follow-up call to the cardholder.

"The service level agreements we have with this bank require us to send that signal back in less than 4/10ths of a second," said Barry. "We go through scoring the rules, anomaly detection, predictive modeling and link analysis, and send that signal back in less than 4/10ths of a second. This system does this for 4.5 million trans-actions per day. That gives you an idea of the amount of data we can handle and the speed at which analytical processing on that data can be done."

Governments now have new opportunities to scan and score tax records in ways that have not previously been feasible because the data volumes were just too big.

Tax Fraud Analytics in Action

Case 1: In-Stream Scoring for Her Majesty's Revenue and Customs

The United Kingdom has a significant tax gap – and it is rising, particularly with the economic malaise that is overtaking most of Europe. Her Majesty's Revenue and Customs adopted analytical methods to do in-stream return processing to address the issue. As returns come into the system – either electronically or manually – they are scored for the risk of compliance and evasion. Refunds or rebates are stopped from going out the door if the submission is deemed highly suspicious, enabling the agency to take a proactive step in managing tax fraud.

Scoring is based on multiple years of tax returns, which sounds easy in concept but wasn't practical before, said Lemon. "Only in the past few years has the technology evolved to the point of being able to handle all those massive amounts of data in a very short time. Now HMRC can use all that incoming and historical data in near-real time to get a very complete picture of who's cheating and how they are doing it."

So far the agency estimates about £10.5 billion in savings in just a few years – either refunds/rebates that would have gone out the door but were stopped, or taxes received that otherwise would not have been identified or collected.

Case 2: The Phantom Tax Preparer

Phantom tax preparers are accountants or tax preparers who prepare tax returns for others but don't identify themselves as such on the clients' return. They don't want tax agencies to be able to connect the dots and link them to multiple suspicious returns.

"Our analysis found a tax preparer who had submitted more than 1,900 tax returns in three years, of which more than 95 percent requested a refund for her clients," said Barry. "In none of the returns was there an indication of a tax preparer, but link analysis found that all the returns were electronically submitted from a single IP address in increments of 15 to 30 minutes.

"The real kicker here is that there's nothing suspicious about those factors. What was suspicious is that anomaly detection found very large charitable contributions and job-related expenses on these returns, relative to the taxpayers' peers. The state tax agency might have stopped individual cases as they came through and looked at them, but it probably doesn't have enough staff to go after all 1,900 individual taxpayers. It's much more efficient to pursue one tax preparer who is responsible for so many acts of fraud."

Case 3: Geo-boxing to Find Viral Tax Fraud

Imagine a case where a friend, neighbor or colleague says, "Hey, I just got a big refund check, and I bought a motorcycle over the weekend." Everybody wants to know how he did it. And the response is, "Well, you know, I went in, and I fudged on these numbers, or I made this number look like this." Friends and family members want to benefit as well, and the tactic spreads by word of mouth.

"In this case, we used a technique called 'geo-boxing,'" said Lemon. "It's a technique commonly used by marketing to create microsegments or targeted neighborhoods of people who live in a similar geographic area and have similar types of behaviors. We divided the entire state into neighborhoods of no more than 250 taxpayers to see where people are (or should be) behaving in similar ways, and then looking for anomalies from that geo-box norm."

For example, suppose a taxpayer who had always used a standard deduction suddenly shows a large number of itemized deductions and claims an earned income tax credit. Taken alone, the return might be perfectly reasonable, or it might not. "When we used geo-boxing, we found a fascinating result," said Lemon. "When we looked at other taxpayers in the same geo-boxing neighborhood, we found more than a dozen other taxpayers whose returns showed exactly the same patterns, even almost the exact same adjusted gross income. This was to a signal to us that viral tax fraud was going on. People are saying, 'Hey, how'd you get the money?' And it's being passed along in an informal ring."

Case 4: Is It Legitimate or Is It Identity Theft?

Time series analysis found that a child in Oregon was declared as a dependent of one taxpayer in April one year, of another in March the next year, and a dependent of a third person as soon as electronic filing opened the next year.

"What's interesting about this example is that it could be a perfectly acceptable social situation," said Lemon. "The girl might have lived with her father the first year, her mother the next, and her grandmother the third year. These social circumstances happen all the time. Or, it could be an example of identity theft, where the third person has stolen the girl's Social Security Number and claimed her as a dependent to get a big refund early in the season, before her mother or father submit their tax returns.

"Analytics can help tax administrators discern the difference between these two situations – one perfectly legitimate and the other perfectly fraudulent. Tax administrators don't want to get into the middle of a possible custodial battle. They just want to find out who is reporting tax liability correctly."

> "Analytics has helped tax administrators around the country get much smarter, much more precise in the way that they can go through and tackle and discern very complex situations."
>
> Jon Lemon,
> Solutions Specialist,
> SAS Federal

Closing Thoughts

The successes of tax administrators provide lessons that can be applied in other industries and focus areas. Lemon and Barry shared top takeaways that have broad application outside of tax authorities.

Know your weaknesses. As we saw with the IRS, where trusted third-party information is available, compliance is very high and fraud is remarkably low. Where information reporting is lacking, that's where people are cheating. Governments are moving aggressively to find new data sources that can improve the predictive and descriptive power of their analytics.

There is so much untapped potential, even within government databases, said Lemon. "Where you have government programs, you have data in silos. You can go on and on with the list of data sources out there – worker's compensation data, Medicaid/Medicare program data, social benefit programs such as welfare or food stamps, motor vehicle registrations, courts and corrections. By bringing multiple data sources together, you gain insights that might not have been revealed otherwise. Governments are starting to get pretty smart about selectively sharing data when it makes sense to address tax evasion and other types of fraud."

Take advantage of advanced analytics. "If you had the time, you might be able to find sophisticated fraud schemes just by doing gumshoe detective work or sleuthing," said Lemon. "But advanced analytics enables you to quickly discern those really hard-to-detect schemes, the places where tax authorities are losing the most money."

Think creatively about how other industries use analytics. There are only so many ways to cheat a financial system. While the problems may look different on the surface, patterns of fraud are very common across industries. So look at how government entities, health care insurers, property and casualty insurers, or any type of financial services entities are adapting analytic techniques to address fraud and improper payments.

Don't let the perfect be the enemy of the good. Do you have qualms with the quality and completeness of your data? You're in good company. "Your data doesn't have to be perfect because nobody's data is perfect," said Lemon. "It just has to be good enough to go through and derive some value using these analytical methods. Take what you can out of the analytics and the data that you do have."

> "If you had the time, you might be able to find sophisticated fraud schemes by doing gumshoe detective work or sleuthing. But advanced analytics enables you to quickly discern those really hard-to-detect schemes, the places where tax authorities are losing the most money."
>
> Jon Lemon,
> Solutions Specialist,
> SAS Federal

For More Information

Learn more about the Association of Certified Fraud Examiners or register for upcoming ACFE member programs online at www.acfe.com.

Learn more about SAS solutions for fraud detection and prevention.

To contact your local SAS office, please visit: sas.com/offices