This chapter is about the role of economic analysis in understanding the tax-compliance problem. The carrots and sticks in the title suggest a recalcitrant beast that needs to be seduced into cooperation by an elusive, dangling incentive or beaten into unwilling compliance. The story is, however, somewhat richer than that. The discussion shows that economic theory can play an important part in explaining the underlying mechanisms relating to the economic engagement of the citizen in the funding of public programs and why such engagement—or lack of it—may be an endemic problem for tax administration. It also shows what the natural limitations of the carrot-and-stick analogy may be and what alternative paradigms of compliance could usefully be employed.

The standard microeconomic approach to tax compliance helps us understand the basic schizophrenia that lies at the heart of public economics. This has little to do with like or dislike of government or with approval or disapproval of the way in which the government raises funds or with the mix of goods and services it provides. To appreciate how the standard approach can be useful in designing empirical investigations into tax noncompliance and in formulating policy, it is important to understand what can reasonably be expected from pertinent economic models and what can be expected by way of evidence.

I am grateful for the research collaboration of Ralph Bayer and Carlo Fiorio and for helpful comments from Henry Aaron, John Scholz, Douglas Shackelford, and Joel Slemrod.
No overall modeling framework can be expected to offer an all-encompassing story of the compliance problem, although particular models can provide particular economic insights that illuminate particular aspects of the compliance problem, allowing a piecemeal appreciation of tax administration. Microeconomic models have a further role in showing what may be the consequences of the successful establishment of specific institutions and norms or, indeed, the breakdown of these institutions and norms. Predictions from these models are always conditional upon the appropriateness of the particular institutional set that is assumed.

By definition, evidence is bound to be limited and imperfect; however, data arising from the audit process or from activities collateral to noncompliance activities can be expected to reveal information about subsets of the compliance problem.

I begin by introducing the standard paradigm.

The TAG Model

The taxpayer-as-gambler (TAG) model is perhaps the benchmark economic approach of modeling tax noncompliance.1 It is important to understand the ground rules of this approach and, thereby, its limitations in indicating the way official incentives work on individual decisions to conceal taxable income and, in some cases, to generate taxable income.

The model is based on the elementary choice facing an individual in an atemporal environment. It is nonstrategic, in that no account is taken of the possible conditioning of taxpayer behavior on beliefs about the tax authority’s reaction to its information signals. Government or tax authority actions may be conditioned on personal attributes, but there is not enough information in the system to build in assumptions about best response.

Foundations

Taxpayers are confronted by a classic economic problem of choice under risk. They know the tax legislation, the taxes they are liable for, and the penalty for failing to pay (if they are caught and convicted). Taxpayers also know that the tax authority is not psychic: The authority cannot know their true tax liability unless they report it or unless the authority spends the time and trouble to find

1. The model was pioneered by Allingham and Sandmo (1972) and is widely discussed in the literature. For an introduction see, for example, Andreoni, Erard, and Feinstein (1998, pp. 823–24); Cowell (1990), pp. 50–59; Franzoni (1999, sec. 3.1); Slemrod and Yitzhaki (2002, chap. 4).
out for itself. So taxpayers could get away with concealing part of their resources, falsifying reports made to the tax authority, or even making no report at all. Being without moral scruples, they are tempted to take the opportunity of evasion.

At the heart of the analysis is a simple and familiar lottery: Is it worth the taxpayer’s taking a chance on being caught and suffering a financial penalty? Assume that the taxpayer’s initial resources and all gains and losses can be measured in terms of a single consumption good, which can be termed *income*. To further simplify the discussion, make two important assumptions about time and uncertainty: First, time is compressed into a single period within which the taxpayer has to make a decision on whether to attempt to evade paying tax and, if so, how much to evade. Second, once the decision to evade tax has been taken, one of two possible states occur: Either the taxpayer is not audited and enjoys a consumption level \(c\) or is audited, convicted, and punished, in which case consumption is \(c''\), and \(c'' < c\).

If the taxpayer has chosen not to evade tax, then \(c'' = c\).

The exact nature of the lottery is determined by the taxpayer’s financial resources, the tax system, and the penalty system in force. The model assumes that the tax system is based on income and the following three axioms:

Axiom 1: The individual has a fixed gross income, \(y\), which is liable to tax.

Axiom 2: There is a proportional income tax at rate \(t\).

Axiom 3: There is a fixed probability, \(p\), of tax evasion being discovered and punished, and the tax on any concealed income is subject to surcharge at a rate \(s\).

Viewed this way, noncompliance is just another risky activity with a known distribution of returns. The rate of return, \(r\), to a dollar of evaded tax takes the value \(-s\) with probability \(p\) and the value 1 with probability \(1 - p\). So the expected rate of return is \(r = 1 - p - ps\).

But if the taxpayer has behaved honestly and declared \(y\), disposable income would be \((1 - t)y\). Otherwise, disposable income would depend on the amount of evasion: If an amount of income \(e\) is concealed (so that the taxpayer reports \(y - e\)), then consumption is given by the random variable:

\[
(1) \quad c = (1 - t)y + ret.
\]

An analysis of the taxpayer’s optimal evasion decision, given the above budget constraint, shows that the exact decision that the taxpayer makes depends on his or her personal attributes, \(a\), which may include willingness to take risks and innate honesty. However, it is conventional to assume that all taxpayers of whatever \(a\) type have the same structure to their preferences over the state-contingent consumption levels \(c\) and \(c''\). The standard assumption is

Axiom 4: Each \(a\)-type taxpayer’s preference is represented by an expected utility function:
Where $u^a$ is an increasing, concave function.

What this means is that the utility derived from disposable income (consumption) is increasing but that marginal utility increases at a decreasing rate. It rules out the phenomenon of the risk lover: Everyone is assumed to be either risk averse (in which case the indifference curves are strictly convex to the origin) or risk neutral. Furthermore the slope of any indifference curve in the neighborhood of perfectly honest behavior is fixed at $-(1-p)/p$ (the betting odds on the taxpayer succeeding in his evasion), irrespective of income.\(^2\)

With $c$ determined by the rate of return to evasion $r$ and by equation 1, if the taxpayer conceals some but not all income, then

\[ (3) \quad 1 - p \frac{u^c(c')}{u^c(c')} = s \]

where $u^c(c')$, $u^c(c'')$, denotes the $a$ type's marginal utility in the two possible cases (not-caught and caught), respectively. The simple interpretation is

Marginal rate of substitution = proportional penalty.

In principle, one should also consider two special cases that modify this conclusion. If the taxpayer reports completely honestly, then,

Marginal rate of substitution $\leq$ proportional penalty.

If the taxpayer completely evades, then,

Marginal rate of substitution $\geq$ proportional penalty.

Equation 3 can be used to derive optimal evasion $e^*$ as a function of the tax enforcement parameters $(p, s, t)$ and the personal characteristics $(y, a)$. The properties of this function are inherited directly from the assumptions about the utility function, and it can be used to derive a number of specific behavioral predictions.

The aggregate. If the economy is large, the government may take as deterministic the total amount of revenue it receives through the penalties imposed on proven tax evaders, although the amount each taxpayer has to pay (tax plus surcharges) is random.

The appropriate budgetary constraint upon government might be modeled several ways. The standard version is as follows. The government has a specific net revenue target, $R_0$, and it faces an aggregate resource cost of enforcement that is increasing in the detection probability, $p$. Actual revenue raised, $R$, is

\[ R = R_0 - CR^e \]

2. This the point at which $c'' = c'$, because $e = 0$. 

given by total legal tax burden minus the total leakage through evasion and the resource cost of enforcement. The constraint that the government faces is simply \( R \geq R_0 \): Tax receipts, net of any leakages to the underground economy and administration costs, must be at least as great as revenue. Given an appropriate objective function and a specification of the resource-cost function, this constraint can form the basis for the design of an optimal tax enforcement policy. But uncritical application of this apparently commonsense criterion in a normative model can lead to unfortunate prescriptions.

Results

The following is a brief review of what can be deduced immediately from the basic assumptions of the simple TAG model and the attempts to implement it empirically.

Theoretical Overview. Although three possible outcomes of individual optimization (equation 3 and the two modifications that follow it) are described, only two are relevant. Given that the taxpayer is an expected utility maximizer and that the marginal utility of consumption is positive, then the first outcome drops out if \( r/H_2 > 0 \), that is, if the taxpayer always conceals some income.\(^3\) Equation 3 also shows that increasing the probability of detection, \( p \), or the surcharge, \( s \), will shift the equilibrium in such a way that \( e^* \) is reduced.\(^4\)

Furthermore, there is an intuitively reasonable result to be obtained that characterizes taxpayer behavior across different attribute classes of taxpayers. An \( a \) type’s risk aversion is defined to be the proportion rate at which the \( a \) type’s marginal utility falls with consumption:

\[
\frac{u_a^e(c)}{u^e_a(c)}
\]

\(^3\) See Cowell (1990), p. 56, for an explanation. However, this may not apply in richer models of taxpayer choice. If, for example, the taxpayer problem is temporary and he or she faces borrowing constraints, then the taxpayer may evade the tax even if the expected return is negative. Andreoni (1992). Of course if the taxpayer’s preferences in the face of risk do not conform to that of the expected utility model, then the taxpayer may comply more than the conventional theory would suggest. Bernasconi (1998).

\(^4\) Equation 3 can be rewritten as

\[
\frac{U_a((1 - ty + et))}{U_a((1 - ty + set))} = \frac{pt}{1 - p}
\]

Check the left-hand side of this equation: remembering that \( U^e_a(\cdot) \) is everywhere decreasing or constant we can see that an increase in \( e \) will decrease the numerator and increase the denominator; the LHS is decreasing in \( e \). Now increasing \( p \) or \( s \) obviously increases the right-hand side. So the only way the equation can still hold if \( p \) or \( s \) increases is if \( e \) falls.
where \( \mu_a (c) \) is the slope of the marginal utility function (negative in the case of risk aversion). If a taxpayer’s risk aversion for all values of \( c \) is dependent on personal attributes, then certain taxpayers will always conceal more income than other taxpayers.\(^5\)

To obtain other results, a fifth axiom is usually introduced:

—Axiom 5: Absolute risk aversion is a decreasing function of \( c \). This implies that a risk-averse individual who holds a portfolio containing a safe asset and a risky asset will increase the holding of the risky asset were the endowment to increase. So for any particular \( a \) type and any given set of tax enforcement parameters \((p, s, t)\), if the individual’s taxable income, \( y \), increases, then so too does \( e \), the absolute amount of income concealed. Decreasing absolute risk aversion does not permit anything definite to be said, however, about the proportion of taxable income that is being concealed.

So the elementary analysis of behavior in the face of risk results in four simple propositions about the incidence of tax evasion in the community (also see appendix 8A):

—If the rate of return to evasion is positive, everyone evades tax.
—Those with higher risk aversion tend to evade less.
—Those with higher personal income tend to evade more.
—Increasing any of the tax enforcement parameters \((p, s, t)\) will reduce the amount of concealed income.

The TAG model is remarkably robust, in that the above propositions are established for a wide class of individual preferences. However, only the second proposition and two-thirds of the fourth proposition seem to chime with common sense. It seems strange to assert that all taxpayers will evade; and although one would expect compliance to increase with the probability of audit, \( p \), and with the size of the surcharge, \( s \), why should it also increase with the nominal tax rate, \( t \)? Many would argue that common sense suggests the opposite. As for the third proposition, who knows? Clearly this is an area where common sense is not adequate, and we need empirical evidence.

**Empirical model specification.** The TAG model gives aggregate evasion, which suggests that an appropriate econometric version of the model ought to have tax and enforcement parameters, personal income, and indicators of type of income recipient as explanatory variables; the dependent variable would be some measure of underreported income. The model could be estimated for different categories of taxpayer or for taxpayers in general. The empirical model could be used to test the empirical validity some of the propositions on the shape of the \( e \) function raised earlier.

---

5. See Cowell (1990), pp. 57, 58.
A number of difficulties arise, however, with appropriately specifying an empirical model. There may be underlying problems of sample-selection bias; for example, even a carefully conducted review of taxpayer audits may nonetheless exclude individuals who do not file a tax return at all. Furthermore a particularly tricky difficulty is the specification of the variable characterizing the probability of audit. Usually some proxy for evasion opportunity (such as the presence of business income) has been used to categorize audit classes, and the probability of audit can be expected to differ across these classes. Finally there is a “rationing” problem: Individuals’ opportunities for participating in evasion differ greatly among occupations and social groups, although one might suppose that the membership of rationed and nonrationed groups is largely self-selected. The appropriate margin of choice for an individual may not be to change the amount of evasion undertaken within the context of a particular group but rather to migrate among groups in response to changes in tax enforcement parameters.

**Taxpayer Audits.** In the United States the Tax Compliance Measurement Program (TCMP), a program unmatched in other countries, provides the empirical researcher a preeminent data source. Here I briefly summarize some of the empirical work based on taxpayer audit data, principally from the TCMP.

- Tax compliance differs according to income type and socioeconomic group. For example, it is lower for married people than for single persons, lower for younger people than for older people.
- Source of income rather than income level is a significant determinant of evasion. A much higher proportion of wage and salary income than self-employment income is reported. Those paying taxes on non-farm business place a lower value (about 0.3) on the income elasticity of underreporting than those paying taxes on farm business income (about 0.65).
- Income level and enforcement parameters generally have the expected effects on evasion behavior. Higher income is associated with higher amounts of underreported income. Although taxpayer compliance is usually positively associated with the probability and severity of criminal penalties, the relationship is weak. This is broadly confirmed by a Minnesota experiment, in which taxpayers were informed that they were likely to be subject to close scrutiny (high-

---

6. See, for example, the approach in Andreoni, Erard, and Feinstein (1998).
7. The last TCMP was done in 1988, although a limited version was introduced in 2002 as the National Research Program.
8. The results on taxpayer characteristics and income sources are drawn from the studies by Beron, Tauchen, and Witte (1992); Clotfelter (1983); and Feinstein (1991). The main sources for the role of income and the impact of enforcement parameters are Beron, Tauchen, and Witte (1992); Landsman, Shackelford, and Yetman (2002); Poterba (1987); and Witte and Woodbury (1985). For an authoritative account of the inherent econometric problems, see Feinstein (1991). The Minnesota audit experiment is reported in Slemrod, Blumenthal, and Christian (2001).
income taxpayers appeared, however, to behave differently from low- and middle-income taxpayers).

—Early studies suggest less taxpayer compliance in groups with higher marginal tax rates. However, in any sample of taxpayers taken at any particular time, those facing different marginal tax rates may belong to groups that have different economic opportunities or that have significantly different preferences for risk and attitudes toward evasion. More recent work shows greater insight on the important relationship between the marginal tax rate and evasion, separate from income: While income has only a weak effect, the marginal tax rate has a negative impact.

—Detection is imperfect: Variation in detection rates is at least as important as variations in personal characteristics.

**Indirect evidence.** The microdata on tax evasion are, perhaps understandably, limited in availability and coverage, so some researchers infer underground economic activity by using an indirect indicator, including monetary variables or apparent differences between income and expenditure at the aggregate level.

Unfortunately many of the more ambitious attempts to obtain indirect evidence are of dubious value, since they are only sketchily based on economic theory and suffer from severe econometric shortcomings. What is needed is a careful empirical model of the relationship between observables that appropriately takes account of the influence of unobservables in its specification and that provides a plausible basis for distinguishing between the impact of noncompliance and other unobservables.9

**Laboratory experiments.** The questions that either microdata or indirect evidence allow the researcher to pursue usually concern issues such as the possible role of specific personal or job characteristics as factors predisposing to tax compliance, the impact of changes in tax rates, and the relationship between compliance and the tax structure. It is not usually possible to focus clearly on taxpayer motivation, which may be of immediate concern for those who want to judge the effects of incentives—the carrot and the stick—on compliance.

Experimental methods suggest themselves as a possible way of filling this gap. It is not often that circumstances permit experiments with real taxpayers, so it is not surprising that several economists and psychologists have used laboratory methods.9

9. An example of the aggregate approach is Crane and Nourzad (1986), who use a synthetic series of an “adjusted gross income gap” as a measure of tax evasion. Modeling this as a function of tax rates, enforcement parameters, income, and the inflation rate, they suggest that aggregate evasion falls with the tax rate, in line with the early cross-section TCMP evidence. Thomas (1999) provides a good overview of the methodological pitfalls in many of these aggregate approaches. A good example of appropriate micromodeling of behavioral relationships is Pissarides and Weber (1989), who use the Family Expenditure Survey (United Kingdom) to model the differential relationship between expenditure and income for the employed (with very low evasion opportunities) and for the self-employed, who clearly have substantial opportunity for noncompliance.
experiments. The results are not encouraging for the TAG model: Early studies conclude that subjects do not act like gamblers in the tax compliance setting, and it is not even clear that they act in conformity with the basic economic model of risk taking. Furthermore it appears that the structure of taxation is important, over and above the levels of tax rates and exemptions. However, the evidence on the responses to tax enforcement parameters is broadly in line with what gets from econometric analysis of the microdata. A higher probability of audit is associated with greater compliance (although it may have its principal impact on whether one chooses to evade at all rather than on the amount evaded) conditional on noncompliant behavior.10

**WHAT IS WRONG WITH THE MODEL?** Some aspects of the TAG model are distinctly unsatisfactory, especially the implication in axiom 5 that, as long as the expected rate of return to evasion is positive, everyone will conceal some income. The structure of the model might be reformed in three areas:

—The nature of taxpayer motivation. The assumption is usually made that the objective function should be in the form of expected utility. But what is a more appropriate characterization of risk?

—The nature of the economic interaction. Because of the inherently non-strategic nature of compliance, some essential features determining compliance and the possibility of manipulation by the tax authority may have been assumed away. Furthermore the atemporal setting arguably leaves out some of the crucial aspects of the interaction between taxpayer and tax authority (for example, it completely misses the issues associated with tax amnesties).

—The nature of the economic agent. The model assumes gamblers endowed with exogenously fixed incomes. While this assumption has been relaxed in some models to include labor supply, the productive economy is usually ignored.11 In particular, focusing on the TAG model typically neglects a key feature of tax noncompliance: the behavior of firms. Given that the firm is constrained only by the size of the market and its ability to undercut the costs of competitors, the consequences of successful individual attempts at noncompliance may be enormous.

**Rethinking Taxpayer Motivation**

Underlying the TAG model is the simple greed assumption conventionally made in economics, but this assumption may tell only a partial story about the


11. See, for example, Baldry (1979); Cowell (1985); Pencavel (1979).
relationship between citizens and the state. The question as to what motivates taxpayers deserves to be addressed. For example, some people may pay taxes and refrain from evasion out of civic duty. In reviewing taxpayer motivation we may Here I examine a number of issues that affect the model.

Expected Utility

The TAG model, rooted in conservative economic theory, assumes rational individuals with stable preferences who, given specific economic opportunities and probabilities, maximize their expected utility. The expected utility (EU) paradigm may be good as a device for simplified model building, but it may miss important nuances about people’s preferences in the face of uncertainty. Indeed the use of EU assumptions to characterize these preferences is arguably restrictive. It rules out state-dependent utility and hence any feeling of shame (or delight) at successful evasion. It also rules out regrets and misperceptions on the part of the taxpayer about the probabilities of alternative possible states of the world—for example, the probability of audit. However, there is evidence that individuals make systematic mistakes when attempting to maximize their expected utility. Would relaxing the expected utility assumption to consider other models—such as rank-dependent utility or prospect theory—result in a more promising underlying story?

Rank-dependent utility is unlikely to be a fruitful approach in the present context given the typically uncomplicated nature of the risk involved: The possible outcomes are usually taken to be the simple pair (income-if-not-caught, income-if-caught) rather than a structure of possible payoffs. However, prospect theory incorporates features that may be relevant to the problem of appropriately modeling taxpayer choice, in particular the following:

—Individuals “edit” information about gambles before they evaluate them so as to simplify the representation of the prospect with which they are faced.
—Individuals use a reference point from which to measure outcomes in terms of changes.
—The value function is defined over gains and losses relative to the reference point rather than to absolute values of wealth or income.
—In evaluating gambles, individuals assign decision weights different from the actual probabilities.

The first three of these features lead to a version of the framing phenomenon, in which risk choices are evaluated according to the way they are presented to the decisionmaker; in particular, gains may be evaluated quite differently from losses, relative to a particular reference point.

Several studies that examine tax evasion in the light of prospect theory suggest support for the framing hypothesis, whereby the response to a particular
economic incentive (carrot or stick) differs according to the context. But direct
tests of conformity of behavior with prospect theory are inconclusive.\textsuperscript{12} It is not
clear that prospect theory receives overwhelming support in comparison with
expected utility theory, although the framing issue remains important for the
issue of the effectiveness of incentives and sanctions.

\textit{Range of Goods}

The range of goods in the utility function is also simplified in the TAG model.
Individuals are concerned only with their own private consumption and so of
course care nothing for the goods and services produced by the resources raised
through the tax system. This issue is relevant to economies with a small public
sector as well as to those that supply a lot of goods publicly. One would expect
to see a positive relationship between marginal tax rates and the overall size of
the underground economy if, on average, public goods were perceived to be
underprovided, with the reverse effect if there is overprovision of public goods.\textsuperscript{13}

\textit{The Temporal Model}

The TAG model ignores time. It assumes that, each year, essentially the same
gamble takes place, without there being any “memory” in the system. Some
contributions to the literature attempt to correct this by allowing the tax autho-
ритy to use information from multiple time periods. Even if the tax authority uses
only information from the current period for an audit, the outcome of the audit
may be used to trigger retrospective investigation.\textsuperscript{14} This clearly weights the
“stick” wielded by the tax authority. A rational taxpayer’s current tax evasion is
a decreasing function of evasion in previous periods. The reason for this is that
if the taxpayer is audited and caught evading this year, penalties for earlier non-
compliance may be incurred.

\textsuperscript{12} On the general issues of prospect theory, see Kahneman and Tversky (1979); Tversky and
Kahneman (1981); and Schepanski and Shearer (1995). For a review of experimental evidence on the
expected utility model and other paradigms of individual decisionmaking, see Camerer (1995). Sup-
port for the framing hypothesis is found in the studies by Chang, Nichols, and Schultz (1987);
Robben and others (1990); and Schepanski and Shearer (1995). King and Shefrin (2002) investi-
gated whether individual behavior conforms to the standard results of prospect theory, given a sce-
nario that incorporates a perception of inequity, using a questionnaire with student respondents. The
responses to the control questions are consistent with expected utility theory in that they do not dis-
play the phenomenon of loss aversion (risk taking when faced with losses and risk averse when faced
with gains) characteristic of prospect theory.

\textsuperscript{13} Note that this phenomenon is, nonetheless, consistent with the free-rider problem associated
with public good provision; it follows from the impact of income levels on risk-taking behaviors associ-

\textsuperscript{14} This is the argument in Engel and Hines (1999). See also Greenberg (1984) and Landsberger
and Meilijson (1982) for a detailed analysis of the role of time in the audit sequence.
Interdependence

Among the aspects of taxpayer interdependence that can affect the overall compliance problem is trust and the concept of a climate of behavior. This climate can be represented by an externality in the individual utility function—people may care about their own behavior relative to those of their peers; the stigma or the nonmonetary penalty associated with discovered evasion may be endogenously determined by the behavior of others. This consumption externality may be supplemented by a production externality; the growth in individual noncompliance may facilitate the development of a kind of infrastructure of noncompliance—finding a corrupt accountant for one’s own tax cheating will lower the search costs of other potential noncompliers.

The endogeneity of interdependence in the economic model is crucial. One of the contributions of the economic model that incorporates such interdependence is its explanation of epidemics of noncompliance. The maintenance of a culture of compliance is one example of the government or tax administration creating a “carrot”—a positive incentive for taxpayers to act in their broad social interest rather than in their narrow self-interest.

Donkeys, Mice, and Ghosts

Are taxpayers donkeys? The carrot-and-stick approach to modeling the interaction between taxpayer and tax authority assumes that the tax authority views economic incentives in a fairly simplistic fashion. The reason for this is the simple nature of the economic interaction in the basic TAG model: The tax authority lays down ground rules for the mass of taxpayers; each taxpayer then assumes that the probabilities in the fundamental gamble are uninfluenced by his or her own actions.

Strategic Models

An alternative view of economic interaction between the two parties sees them as cat and mouse, or cat and dog. Each party is aware of the other’s motivations and interests (the taxpayer wants to maximize utility, the tax authority to maximize net tax revenue) and takes these into account in selecting its own strategy. The outcome is an equilibrium in which each party makes the best response to the other’s strategy in the light of the available information.

15. On trust, see Scholz and Lubell (1998b). The stigma model is attributable to Benjamini and Maital (1985). Note that it only requires an aggregate level of externality to be generated by the taxpayers as distinct from the near-neighbor model of Glaeser, Sacerdote, and Scheinkman (1996), in which the position of other economic agents is important.
This approach resolves into two classes of model according to whether one assumes that it is reasonable for the tax authority to precommit to an audit strategy, that is, to set the agenda for the interaction. Which model is the more appropriate depends on factors such as the type of institutions and laws present in the economy and the nature of the information available to the parties.

PRECOMMITMENT. In a model characterized by a simplified distribution of income (just rich and poor in known proportions) and in which the tax authority moves first strategically, the optimal policy of the tax authority is stark. It should audit all low-income tax reports and ignore all high-income tax reports. However, under such circumstances no high-income person would ever dare to report a low income; so in fact the only people who would ever get audited are those who are genuinely low income! The statement of the model may seem extreme, but it contains an inner truth about the regressive nature of such carefully tailored audit schemes.

NO PRECOMMITMENT. By contrast consider a model in which precommitment to such an extreme policy is not credible. Again there are two income levels, but the personal characteristics of the population are such that some will always report truthfully and others will not if they have the opportunity and find it profitable. Let the probability that the tax authority decides to audit a particular low-income report be \( p \) and the probability of a potentially dishonest taxpayer not complying be \( q \). Each party takes fully into account the other's strategy in this game of noncompliance and investigation. The outcome will be a Nash equilibrium characterized by a pair \((p^*, q^*)\) representing the “best response” of each party (the tax authority and the taxpayer) to the other’s strategy. Each of these equilibrium values depends on, among other things, the tax rate, \( t \), the penalty surcharge, \( s \), and the cost of an individual audit. The following generalizations can be made:

—Decreasing the marginal cost of audit (that is, making the investigation and enforcement system more efficient) reduces the probability of noncompliance, \( q^* \), but leaves the optimal probability of audit, \( p^* \), unchanged.

—Increasing the surcharge, \( s \), reduces both the optimal probability of audit and the optimal probability of noncompliance. The first of these is attributable to the usual marginal deterrent effect of higher punishment. The second emerges because the tax authority does not need to put in so much effort to achieve a given net revenue.

—Increasing the tax rate reduces the probability of noncompliance, \( q^* \), and will either increase or leave unchanged the optimal probability of audit.

16. See Reinganum and Wilde (1985) for the model with precommitment and Graetz, Reinganum, and Wilde (1986) and Reinganum and Wilde (1986) for the no-commitment case. The specific no-commitment model discussed here is briefly outlined in the appendix to this chapter.

17. A word of caution: This argument about the marginal deterrent effect in this and other models cannot be pressed too far. See remarks under “Guidance for Policymakers,” below.
Note that, despite the different premises of this model, the impact of the key parameters $s$ and $t$ on compliance is in the same direction as in the TAG model.

**Ghosts**

Ghosts are individuals who fail to comply with their income tax filing requirements in an extreme form: They disappear from the system. From the point of view of economic modeling there is an essential difference between those who make a zero-income report and those who make no report at all. What do we know, or what could be known, about ghosts and the way they can be expected to respond to economic incentives designed by the tax authority?

Unfortunately, information about the behavior and characteristics of ghosts is sparse, although enough is known to suggest that they are quantitatively important: The U.S. ghost population in 1988 is estimated at nearly 8 million (compared with 110 million who filed tax returns); the tax shortfall for ghosts is estimated at $11 billion, or some 15 percent of the known tax shortfall of those who file returns.\(^{18}\)

**A Hybrid Model**

The role of ghosts in the tax enforcement story may, however, be more important than their numbers suggest. Typically, both ghosts and strategic players are present in the same population. The margin between the two types of behavior may be crucial from the point of view of policy design: An overzealous approach to enforcement in the sector populated by strategic players may encourage them to migrate to the ghost sector, which is, in essence, nonstrategic and in which the costs of detection and enforcement are typically higher.\(^{19}\)

**Firms**

Why consider corporations and businesses separate from individual taxpayers in economic models of tax compliance? Some theorists have adopted an essentially pragmatic approach, arguing that to distinguish corporate and personal sectors is a way to understand the overall distortional impact of tax evasion.\(^{20}\) However,

---

18. The results come from Erard and Ho (2001), who extend the standard TAG model to account for nonfilers and use a special subset of the TCMP data containing detailed tax and audit information for both filers and nonfilers of U.S. federal income tax returns.
this is distinct from the issue of whether the underlying economic analysis of tax evasion is, or should be, different according to the sector considered.

A brief look at the economics literature on compliance and tax enforcement as it relates to the behavior of firms reveals the key issues that might characterize a theory of compliance by firms. This theoretical approach could then form the basis of appropriate empirical models for the corporate sector and enable policymakers to develop a quantitative model for analyzing the effectiveness of tax compliance regimes.

In principle, firms can evade by misreporting or making false declaration about profits, sales, or input use and other costs. Does the assumed market environment of the firm make a difference to its compliance behavior?

A Simplified Model

Let us take a simplified model of a firm with constant average and marginal cost, producing a single output subject to tax at a uniform rate. In the same way as described earlier, and so again there is an implied expected rate of return to noncompliance $\bar{r}$. Checking through the formal specification of this model in appendix 8A, it can be seen that this could be reinterpreted as a model in which profit is the tax base. The firm has two types of decisions to make: the quantity of output and the extent to which it conceals output or hides profit.

In analyzing the solution to this double problem, it is useful to introduce two new concepts. The first is the expected tax rate on output $\bar{\bar{t}}$, which is given by the nominal tax rate, $t$, multiplied by a factor of 1, minus the proportion of output concealed times $\bar{r}$. The expected tax rate is under the control of the individual firm (through the choice it makes on concealment) as well as the tax authority. The second concept is the average concealment cost per unit of output, $g$, which is a function of the amount of concealment (amount of tax evasion) undertaken.

Assume that the firm chooses the output and level of evasion to maximize expected profits. Because the model is so simple, expected profits can be written as

$$(P - m - g - \bar{\bar{t}}) \times \text{output},$$

21. The simplified model is based on the standard approach in the literature. The main references are Cremer and Gahvari (1993) and Virmani (1989), who focus on a competitive industry; Marrelli (1984) and Marrelli and Martina (1988), who deal with noncompetitive firms that are assumed to be risk averse; and Myles (1995), who assumes risk neutrality. For a detailed treatment, see Bayer, Cowell, and Fioro (2003).
where $P$ is the price, $m$ is marginal cost, and the components $g$ and $t$ (but not others) depend on the amount of concealment. A number of conclusions immediately follow:

—If the firm conceals output, it will do so up to the point where the marginal cost of concealment equals the marginal reduction of expected tax rate.

—The firm will always conceal some output if $t > c$. This is equivalent to the requirement that the expected rate of return $r$ be positive.

—There is a fundamental separability property between the concealment decision and the output decision. Here the concealment decision is independent of the output decision.

—Output decisions for the competitive firm are determined by a modified “price = marginal cost” rule.

The solution to the maximization problem can be used to derive comparative statistics results in the usual way. In the case of the competitive model we then find:

—Reported sales decrease as the tax rate increases.

—An increase in tax increases the price but by less than the amount of the tax, since some of the tax increase is absorbed in increased evasion.

—An increased probability of detection, $p$, or an increased surcharge, $s$, will raise the proportion of sales declared, expected tax, and the market price.

So, as in the TAG model, enhanced deterrence will have the appropriate effect on evasion; in addition it moves expected taxes in the direction that we might have anticipated. But in contrast to the TAG model, there is an unambiguous prediction of a rise in tax evasion with a rise in the tax rate.

Moreover the results are not special to the competitive model. Under risk neutrality the separability property holds, and so it is not surprising to find that basically the same conclusions apply to the monopolistic case as those for the case of perfect competition. The only real difference in the equilibrium is that the “price = adjusted marginal cost” rule is replaced by a condition involving the elasticity of demand.

However, the separability issue is potentially more problematic once one drops the assumption of risk neutrality. This matters both because it clarifies the factors that determine equilibrium compliance by firms in a variety of market environments and because it allows for clearcut conclusions about the impact of policy parameters.22

22. Wang and Conant (1988) study the expected utility function when a monopolist overstates production costs in order to reduce taxable profits. The uncertain monopolist’s optimal rate of output is not affected by either the profit tax or the penalty rate. Yaniv (1995)’s model of tax evasion covers different types of taxes that can be evaded by the firm showing that these taxes do not alter the separability conclusion. Yaniv (1996) extends the analysis of separability to cases in which both the probability of detection and the penalty rate vary with the amount of cost overstatement. Lee (1998)
An Empirical Analysis

Unfortunately the empirical analysis of corporate tax evasion is extremely limited. In the main it consists either of a compilation of rather obvious results (for example, tax evasion depends on the preferences of the person who has the power over declaration) or of procedures that could be considered as methodological very weak. The main reasons are the lack of theoretical models, since theory mainly focuses on personal income tax evasion; the lack of corporate income tax compliance microdata; and the lack of confidence in microdata on tax compliance and relevance of measurement error.

However, some evidence is available, again drawn from the TCMP. Of special interest are two main results that have no counterpart in the literature on personal income tax compliance:

—A firm’s compliance is positively associated with being publicly traded and with belonging to a highly regulated industry.

—Having low profits relative to the industry median is correlated with higher corporate tax evasion.

Clearly both of these findings have potentially important implications for the design of policy.

An Assessment

What makes the simple microeconomic model of the firm essentially different from the TAG model as it is applied to the individual? Three features stand out: the nature of the taxpayer, the assumption about risk preferences, and the determinants of responsiveness to economic incentives.

The taxpayer. It is reasonable to argue that individuals—and perhaps even families and households—exist as exogenously given entities; the set of potential taxpayers could be imagined as exogenously given. This is not the case with firms. Firms are born and dissolved, they merge and change their shape, and they do all this in response to economic incentives. The tax system and its enforcement mechanism are essential components of those economic incentives, and so a reasonable model of firm behavior has to be established before the impact of tax and enforcement policy on firms can be understood. Of course the

shows that the separability property and the neutrality of profit taxes depend on the audit probability and penalty rate are formulated. See also Wang (1990).

23. Rice (1992) uses a corporate subset of TCMP based on an examination of the tax and financial records of a stratified random sample of about 30,000 U.S. corporations out of a total of 1.5 million corporations with assets less than $10 million.
contrast with the household sector is somewhat overstated, but this contrast may contain an important component of the problem.  

RISK PREFERENCES. A major feature of the model of personal noncompliance is the role of risk aversion in the equilibrium. Although several studies using the cat-and-mouse tradition of strategic models assume risk neutrality, a reasonable amount of risk aversion is required in the TAG model in order to get interesting answers. Risk neutrality or extremely high risk aversion would always yield a solution in which the individual either evades all the tax or none at all.

By contrast, in modeling noncompliance among firms it is common to assume risk neutrality. An assumption of neutrality is often important for the separation result, which allows predictions to be established from the theory.

RESPONSIVENESS TO INCENTIVES. Following on from the difference in conventional assumptions about risk preference is the question of what drives the taxpayers' responsiveness (or lack of it) to sticks and carrots. In the case of individual taxpayers, differences in risk preference characterize the differential: the differing responsiveness of different groups to penalties or to the probability of detection. In the case of the firm, it is quite different. Equilibrium is determined by a first-order condition involving the marginal concealment cost and the rate of return to tax evasion. This essential difference gives rise to, among other things, the different relationship between compliance and the tax rate in the models of the personal and the corporate sector.

Understanding the nature and the determinants of the cost-of-concealment function, $g$, is essential to understanding what is going on in firms' noncompliance and to understanding the economic incentives that may usefully be applied by a tax enforcement agency. The academic literature on this point is rather sparse, but one could conjecture that the following are key factors:

—The nature of the product. The output or sale of a highly visible physical good is harder to conceal than some services, for example. Just as the opportunities for evasion in the individual sector differ strongly across occupational categories (employment versus self-employment), so also one would expect to find systematic differences across industry categories.

—The size and organizational structure of the firm. Firms with a more complex organization are likely to have higher concealment costs: The more people brought into the plot, the greater the security problem and the greater the risk of discovery.

24. In particular one would expect to find a relationship between the individual personal motivations of those running firms and firm behavior. An interesting example of this is Joulaian (2000), who finds a positive and significant correlation between managers' preferences and firm compliance using U.S. data.
—The role of reputation. Firms with a respected brand name have much to lose by exposure of illegal activity and therefore have high concealment costs.
—Degree of concentration of the industry. There are two counteracting effects. On the one hand, an industry with a large number of firms may be easier to police by an external agency: Those deviating from the norm in terms of reporting will be easier to spot, leading to higher concealment costs for each firm. On the other hand, the presence of a large number of similar firms could encourage the spread of concealment technology among them.

Guidance for Policymakers

Unsurprisingly, the appropriate guidance for policymakers depends on the specific model considered. Alternative models can yield useful policy lessons.

*The TAG Model*

If noncompliant taxpayers are, in economic terms, indistinguishable from gamblers, then they should be responsive to the same kind of economic incentives as are gamblers. There may be enough information about individual types, $a$, to tailor an audit policy conditioned on personal characteristics. Of course, such a policy would have to use proxies for the true values of the components of $a$, which are unobservable, but reasonable proxies may well be available. Obviously this approach rests on the assumptions that the TAG model is appropriate, that individuals’ perceptions of the gambles involved are accurate, and that these individuals are rationally pursuing a policy of ex ante utility maximization. But as we have seen, the evidence on this is not particularly convincing.

However, let us take the TAG model at face value for a moment: What recommendations does it suggest? Suppose the objective of the tax authority is simply to raise revenue. Intuition would then suggest (and formal analysis confirms; see appendix 8A) that enforcement should be intensified until the probability of audit satisfies that the marginal revenue raised equals the marginal resource cost.

Marginal revenue raised includes both direct revenue (tax uncovered plus surcharge) and the indirect revenue yielded by the effect on compliance of a higher audit probability. Allowing for the problem that the computation of marginal revenue raised relies on taxpayer perceptions of probabilities that may be inappropriate, the above condition seems to have a commonsense appeal. Moreover, this marginalist rule can be adapted and extended to other versions of the objective function.

There is a snag, though, a result of focusing on just the audit probability as a policy variable. If parameter $s$ is also chosen, then it appears as though the tax
authority can do better; that is to say, it can achieve its objective at lower resource cost by raising \( s \) and cutting \( p \) (as long as \( p \) remains positive). Indeed, given this greater flexibility, there is an obvious method of guaranteeing total compliance: choosing \( s \) and \( p \) to ensure that the expected rate of return to compliance is not positive. If one presses the simple logic of the TAG model, then it is clear that the tax authority should save resources by using a (very) big stick version of deterrence.\(^{25}\)

But uncritical application of the big stick approach can lead to ridiculous outcomes. One cannot assume that such ridiculous outcomes will be prevented by the common sense of legislators or administrators, or one could reach the extraordinary situation of eighteenth-century London, described by Charles Dickens in *A Tale of Two Cities*:

But indeed, at that time, putting to death was much in vogue with all trades and professions and not least of all with Tellson’s [Bank]. Death is Nature’s remedy for all things, and why not Legislation’s? Accordingly the forger was put to Death; the utterer of a bad note was put to Death; the unlawful opener of a letter was put to Death; the purloiner of forty shillings and sixpence was put to Death; the holder of a horse at Tellson’s door, who made off with it, was put to Death; the coiner of a bad shilling was put to Death . . . not that it did the least good in the way of prevention—it might almost have been worth remarking that the fact was almost exactly the reverse—but it cleared off (as to this world) the trouble of each particular case and left nothing else connected with it to be looked after” [book 2, chap. 1].

Of course it is a cheap shot to thus picture the outcome of the simplified economic model. What is more useful is to identify the economic reasons behind the ridiculous outcome of a high \( s \) and a low \( p \) and, perhaps, the way to derive a more useful model. Four faulty characteristics of the simplified model—unreasonableness, ineffectiveness, inequity, and misspecification—are considered.

**Unreasonableness.** At a first glance the obvious objection to the big-stick approach is that it is just not reasonable. Do we really want to see extreme penalties for minor infringement of the tax law? At the very least, legislators and those implementing the law need a sense of proportion as to what is appropriate in the context of taxation relative to, say, fraud and theft elsewhere in society.

**Ineffectiveness.** The TAG model also ignores the issue of appropriately structured punishment. If for a terrible moment one imagines the death penalty for tax noncompliance, we might well also comment, “not that it did the least

\(^{25}\) On the welfare consequences of the simple TAG approach, see Cowell (1989); Kolm (1973).
good in the way of prevention.” Taxpayers would, with impeccable economic logic, conclude that they might as well be hung for a sheep as for a lamb.

INEQUITY. Ex post inequities are almost bound to occur, but it is the job of a sensible tax administration to make sure that the consequences are not grotesquely magnified. An obvious source of potential inequity are errors by taxpayers and auditors. Although the standard model assumes that noncompliance is a result of optimization amoral on the part of taxpayers who desire the public benefits of the state without paying the private cost, a substantial amount of noncompliance could be attributable to mistakes or the outcome of inertia or laziness.

A more sensible approach to the normative analysis of compliance is to allow that errors are entirely possible; indeed this seems reasonable in the light of the evidence from the psychological literature. Taxpayers can be encouraged by appropriate incentives to take care in reporting, while the design and implementation of the penalty structure can distinguish between minor infractions and serious violations, even if this were to be at the apparent cost of some expected revenue.26

MISSPECIFICATION. Although the TAG model has the advantage of conformity with mainstream economic analysis, and although it may be useful as a useful starting point for discussion among those raised in a neoclassical tradition of applied welfare economics, it takes the wrong direction because it is built of the wrong components. Let us consider what might be learned from some of the alternatives that have been mentioned previously.

**Modified Motivation**

A better understanding of how individuals reach decisions under uncertainty can help in the design or modification of a policy to enhance tax compliance.

**Non-EU Models of Risk.** One of the main reasons for the failure of the TAG model is a popular misperception of the probability of audit. The use of decision weights that differ from the actual audit probabilities may give the tax authority an opportunity to induce greater compliance by exploiting this misperception, since it is in its interest that taxpayers overestimate their chances of being caught.

However, the non-EU risk model also suggests that there could be fruitful and low-cost possibilities for administrative innovation. If we take the framing phenomenon seriously there may be considerable scope for imaginative redesign

26. On the role of inertia, see Smith and Kinsey (1987). Boadway and Sato (2000) examine the effects of unintentional errors on the design of tax enforcement and tax policy. Maximal sanctions are not applied, unlike the Dickensian model. Although intentional evasion can be deterred by carrots, inadvertent tax evaders are not protected by either the carrot or the stick.
of conditional payments associated with tax-enforcement. Even though two payment schemes may be formally equivalent in terms of a taxpayer's conditional budget constraint, they may be viewed differently by the taxpayer making the choice under uncertainty. For example, should the tax authority consider prizes for promptness instead of penalties for late payments? One might even suggest that bonuses for an excellent compliance record may be more effective in some cases than surcharges for underreporting.

Finally, since there is evidence that, contrary to the TAG assumptions, taxpayer perceptions are important, the structure of taxation as well as the magnitude of the incentives should be taken into account.

**Interaction models.** Interaction models pick up on an important externality present in the economic problem of compliance. The message of the “epidemic” model is that the impact of a modification in tax enforcement policy should not be judged just in terms of its marginal impact on the compliance of a representative taxpayer. The tax authority also has a role—if not a duty—in fostering a climate of compliance.

Unfortunately the message is mainly negative. The right climate can be lost through careless implementation more easily than it can be built from scratch through group effects and socially responsible behavior. Insofar as the externality is generated by an infrastructure of noncompliance, it makes sense to regulate the activities and institutions associated with this infrastructure. Other forms of regulation in the economy may be crucial for effective regulation of tax compliance.

**Strategic models and hybrids.** The strategic, or cat-and-mouse, model is informative for the design of enforcement strategy in a reporting context. However, it assumes a well-defined and rather limited set of possible outcomes and a highly simplified distribution of unknowns (for example, in the implementable versions of such models there is usually a simple representation of the income distribution from which the taxpayer is assumed to be drawn). This model of the compliance problem seems more appropriate to the one-on-one negotiation between the tax authority and large taxpayers, personal or corporate, rather than to the masses.

Even where the simple cat-and-mouse model is applicable, the model can lead to some uncomfortable conclusions. Typically the kind of tailored policy that emerges from the model generates a regressive application of the tax law: Reports from the poor are audited much more intensively than those from the rich, but for good economic reasons. But these good economic reasons may not be sufficient to recommend a strategy that could be socially divisive.

Furthermore, the lesson of the hybrid model indicates the possibility of spillover: an induced migration by taxpayers from the reporting sector to the ghost sector. Overzealous enforcement in areas with relatively low-cost information may exacerbate the problems in high-cost areas (where the ghosts are).
The Firm

The simplified model of the firm outlined above can yield optimal tax enforcement rules for noncompliance by the firm.\textsuperscript{27} But the more interesting use of the model is to provide working guidelines for those who design tax compliance schemes. Here concealment costs and their relationship to characteristics of the firm seem crucial and will determine the responsiveness to incentives of all firms for which the expected rate of return to noncompliance is positive. It suggests that the right approach to the empirical modeling of compliance and to the practical enforcement of tax payments by corporations should be piecemeal. The appropriate piecemeal approach will depend on the type of market in which the firm operates, the nature of its products, and the size of the firm itself.

Several points from the analysis of the individual sector play an important part in tax enforcement policy toward firms.

—The time component is, possibly, more important for firms than for individuals. One can expect reputation to be relevant for the effectiveness of enforcement mechanisms.

—As with individuals, audit data on firms are bound to be limited, in that they have relatively little to say on ghosts. For many developed economies an important contribution to the understanding of firms’ noncompliance is a suitable model of the underground economy.

—Sometimes practical economic inquiry has to proceed by stealth. The tax authority needs to identify observables that are correlated with profit (as consumption with income) and that firms have an incentive to reveal.

—This suggests that, as with the control of the infrastructure of personal noncompliance, an appropriate compliance policy will go hand in hand with effective regulation of industry.

A Final Word

Although the standard economic model of the carrot-and-stick approach to tax enforcement is flawed in many ways, it is a useful starting point for understanding the mechanics of individual decisionmaking. But it can be misleading as a guide to policy advice. However, this should not make one skeptical of the contribution that theory can make to tax administration: Careful microeconomic analysis of the role of incentives can reveal a lot, if the model is selected with prudence.

\textsuperscript{27} See the derivations in Cremer and Gahvari (1993) and Etro (1998).
Appendix 8A

Elements of the theoretical models that undergird some of the principal assertions in the text are presented here.

The TAG Model

Given the model in axioms 1–4, the first-order condition for maximizing equation 2 with respect to \( e \) is given by

(A1) \( \epsilon[ru_c(c)] \leq 0 \) if \( e^* = 0 \)

(A2) \( \epsilon[ru_c(c)] \geq 0 \) if \( e^* = y \)

(A3) \( \epsilon[ru_c(c)] = 0 \) otherwise.

where \( u''(c) \) denotes the first derivative of \( u' \) and \( \epsilon \) denotes the expectations operator. Inequalities in equations A1 and A2 represent, respectively, the cases in which the person reports truthfully (conceals no income) and in which the person conceals everything. Equation A3 gives the case in which the person conceals just a part of his or her income from the authorities. First-order conditions (equations A2 and A3) can be solved to yield the taxpayer response function,

(A4) \( e^* = e(\tau, y, a) \)

where \( \tau = (p, s, t) \) is the collection of tax and enforcement parameters.

Comparative Statistics. If the person is risk averse and at an interior equilibrium, then equation A3 characterizes the optimum, and differentiation can be used to obtain the way \( e \) changes in response to policy parameters. For example, differentiating equation A3 with respect to \( p \) and using equation 1, gives

(A5) \( \epsilon[r^2u''(c)] \frac{\partial e(\tau, y, a)}{\partial p} - u''(c') - su''(c'') = 0 \)

The expectation term on the left-hand side must be negative, in view of the concavity of \( u'' \), and so

(A6) \( \frac{\partial e(\tau, y, a)}{\partial p} = \frac{u''(c') + su''(c'')}{\epsilon[r^2u''(c)]} < 0 \)

Likewise, one can derive
and, if axiom 5 holds and if $s$ is a constant independent of $t$ and $y$, then

$$\frac{\partial e(\tau, y, a)}{\partial s} < 0$$

Note that equation A8 holds if the penalty is proportional to the tax evaded (as in my interpretation of the TAG model) rather than to the income concealed.¹

**THE AGGREGATE.** If the number of taxpayers is effectively infinite and the distribution of individuals in the community by $(y,a)$-type is given by a continuous distribution function $F(y,a)$, then aggregate income is

$$Y = \int y \, dF(y, a)$$

aggregate evasion is

$$E = \int e(r, y, a) \, dF(y, a)$$

and revenue raised is

$$R = tY - rE - \Phi(p),$$

where $rE$ is the expected aggregate loss through tax evasion and $\Phi(p)$ is the dollar cost to the government of enforcing the probability of detection $p$ everywhere.

A rule for public policy can then be derived by differentiating with respect to $p$:

$$\frac{\partial R}{\partial p} = -\frac{\partial(E)}{\partial p} - t - \frac{\partial \Phi(p)}{\partial p}.$$

So if the objective were simply to maximize revenue, $R$, setting equation A11 to zero would yield:

$$\left[1 + \frac{e}{t}\right] \left[\frac{\partial e}{\partial p}\right] = \frac{\partial \Phi(p)}{\partial p}$$

**Cat and Mouse**

Consider a world in which there are exactly two levels of income, $y_0$ and $y_0 + \Delta y$, and three groups of taxpayers with characteristics known to be as in the following table:

Consider first the taxpayers’ position. The behavior of those in groups 0 and 1 is fixed; and those in the group 2 get expected utility

\[(A13) \quad pu^c[(1 - t)y_0 + (1 - t - st)\Delta y] + (1 - p)u^c[(1 - t)y_0 + \Delta y] \]

if they cheat and

\[(A14) \quad u^c[(1 - t)(y_0 + \Delta y)] \]

if they do not cheat, where \( p \) is the assumed probability that a low-income report will be audited. The value of \( p \) that equates equations A13 and A14 is given by

\[(A15) \quad p^* = \frac{u^c[(1 - t)y_0 + y] - u^c[(1 - t)(y_0 + \Delta y)]}{u^c[(1 - t)y_0 + y] - u^c[(1 - t)y_0 + (1 - t - st)\Delta y]} \]

If the person were risk neutral, then equation A15 becomes \( p^* = \frac{1}{1 + s} \).

Let \( q \) be the proportion of group 2 who cheat on taxes. If they believe that \( p < p^* \), then all will cheat (\( q = 1 \)). If they believe that \( p > p^* \), then none will cheat (\( q = 0 \)).

Now consider the tax authority. It knows that group 0 has to report \( y_0 \), that the group 1 people feel bound to report \( y_0 + \Delta y \), and that each person in group 2 could report low (\( y_0 \)) or high (\( y_0 + \Delta y \)); it assumes that a proportion, \( q \), of this group will report low. If the authority aims to maximize net revenue and audits a proportion, \( \rho \), of the low-income reports, then the probability of catching an evader is

\[ \frac{f_2q}{f_0 + f_2q} \]

So if the cost of an individual audit is \( \varphi \), the expect net revenue from the policy is

\[(A16) \quad [f_1 + (1 - q)f_2]t\Delta y + \frac{f_2q}{f_0 + f_2q}p(1 + s)t\Delta y - \varphi \rho, \]

which may be rewritten as
From equation A17, if \( q > q^* \), then expected net revenue would increase everywhere with \( p \), in which case the authority would investigate all low-income reports (\( p = 1 \)); but if \( q < q^* \), then expected revenue would decrease with \( p \), and the authority would choose \( p = 0 \).

The Nash equilibrium is given by the point at which the beliefs of the tax authority and those of the taxpayers are consistent. This is the point at which \( p = p^* \) and \( q = q^* \). To see how this equilibrium is affected by public policy, one would differentiate equations A15 and A18 with respect to the parameters \( \varphi, s, t, g \), giving:

\[
\begin{align*}
\frac{\partial p^*}{\partial \varphi} &= 0, \quad \frac{\partial q^*}{\partial \varphi} > 0, \\
\frac{\partial p^*}{\partial t} &\geq 0, \quad \frac{\partial q^*}{\partial t} > 0, \text{ and} \\
\frac{\partial p^*}{\partial t} &< 0, \quad \frac{\partial q^*}{\partial t} < 0.
\end{align*}
\]

**The Firm**

The simplified model uses the following assumptions:

—Proportional cost function: Average and marginal cost are a constant, \( m \).
—Proportional tax: Output, \( x \), is taxed uniformly at rate \( t \).
—Determinate demand: The firm faces a demand function, \( x(P) \), or, equivalently, can command a known price, \( P = P(x) \), for its product, where \( P(\cdot) \) is the inverse demand function. This includes as a special case the situation of perfect competition, where \( P = \text{constant} \).
—Costly concealment: A proportion \( \beta \) of sales are concealed by the firm; that is, a proportion, \( 1 - \beta \), of sales are declared to the tax authority, \( - \), where \( 0 \leq \beta \leq 1 \). The unit cost of concealing is given by \( G(\beta) \), where \( G(\cdot) \) is an increasing convex function.
—Fixed detection probability. The probability of discovery by the tax authority and subsequent conviction is fixed at level \( p \).
Fixed proportional penalty. The penalty rate on evaded tax is $s$.
Hence the expected tax rate per unit of output is
\[ \bar{t} = (1 - \beta + \beta p(1 + s))t = (1 - \bar{r})t \]
where $\bar{r} = 1 - p - ps$, as before. Expected profits are
\[ (A22) \quad [P - m - \beta G(\beta) - \left( (1 - \rho)(1 - \beta)t + \rho(1 + s\beta)t \right) \times (P) = [P - m - g(\beta) - \bar{t}] \times (P) \]
where $G(\beta)$: = $\beta G(\beta)$ is the average concealment costs per unit of output. For any given output level, $x > 0$, equation A22 implies that the firm chooses $\beta$ to minimize concealment costs (as a proportion of total output) plus the expected tax rate, $G(\beta) + \bar{t}$.

The first-order condition for a maximum is
\[ \frac{dg(\beta)}{d\beta} = 0, \quad \frac{\partial \bar{t}}{\partial \beta} > 0, \]
which simplifies to
\[ (A23) \quad \frac{dg(\beta)}{d\beta} = (1 - \rho(1 + s))t \]
\[ (A24) \quad \frac{t - \bar{t}}{\beta} \]
From equation A23 a necessary condition for an interior solution for $\beta$ is that
\[ (A25) \quad 1 - \rho(1 + s) > 0 \]
or, equivalently, $\bar{t} < t$ for $0 < \beta \leq 1$.

Note that equation A25 is exactly the same as the requirement that the expected rate of return to evasion be positive in the simple TAG model. If equation A25 is violated, then clearly no evasion issue will arise, and the firm will report honestly.

Market equilibrium for a competitive firm occurs at $P = m + g + \bar{t}$, implying that expected profits are zero; actual profits are positive if the firm is not audited, negative if audited.

Differentiating equation A23 with respect to $t$, gives
\[ \frac{d^2g(\beta)}{d\beta^2} \frac{\partial \beta}{\partial t} = [1 - \rho(1 + s)] \]
The same method gives

\[ \frac{\partial \beta}{\partial t} > 0 \]

The same method gives

\[ \frac{\partial \beta}{\partial p} < 0, \quad \frac{\partial \beta}{\partial s} < 0. \]

**COMMENT BY**

**John T. Scholz**

Frank Cowell’s analysis provides an insightful, expansive summary of the advantages and limitations of the utility-maximization framework, which underlies much of the current thinking about administrative approaches to tax compliance. My comments underscore some of his main points but primarily expand on one critical limitation of the carrot-and-stick approach—the inattention to the role of justice, or “just deserts,” in maintaining a system of tax compliance. Specifically I argue that the current “crisis” of tax administration rests in part on confusion about the appropriate roles of deterrence and justice in maintaining tax compliance. How can tax administrators integrate the relatively new concerns about taxpayer rights and customer service with the established concerns about audits, sanctions, and criminal prosecutions? Building on Cowell’s argument, I suggest that positive motivations do indeed play a critical role in maintaining taxpayer compliance, but that taxpayers, like most people, prefer just deserts to carrots.

Cowell presents a simple (at least as formal models go) yet sophisticated model of taxpayer-as-gambler and shows how it can be extended to probe a wide range of enforcement and compliance problems relevant to deterrence. He demonstrates that the basic model can be used to derive the assumptions necessary for the standard conclusions of deterrence theory: Compliance increases with the increase of enforcement and sanctioning activities, particularly among risk-averse taxpayers. Of course Cowell also points out that these standard beliefs have only modest empirical support at present. Furthermore he shows that, under reasonably innocuous assumptions, the model predicts that all taxpayers will evade if there is any positive return to evasion, and that higher income is associated with greater evasion, conclusions that he notes are not necessarily intuitively reasonable or empirically supported.

This balance between optimism about the range of tax compliance behavior the model could explain and skepticism about the restrictive assumptions and
limited empirical support for current models is one of the strengths of his chapter. Cowell emphasizes that models based on the taxpayer-as-gambler model can indeed clarify many problems relevant to administrative concerns with maintaining tax compliance, but to do so they generally require less-restrictive assumptions about motivations, greater sensitivity to context, and greater empirical verification to ensure that the model's predictions indeed reflect the reality of the compliance problem being analyzed. Basing policy decisions on untested, overly simplistic models is a recipe for unhappy surprises and short careers. It is easy to support Cowell's call to apply the taxpayer-as-gambler model judiciously and to work harder at empirical verification of the model's implication.

Three sections of the chapter suggest promising directions for further development of the basic model and consider the policy implications of these potential developments. The section on firms contains the best-developed suggestions. By including the cost of concealment as a choice variable and noting that this cost varies systematically across different industries, market sectors, and types of organization, Cowell shows that deterrence effects (and hence deterrence policies) should vary systematically, depending on the type of market in which the firm operates, the nature of its products, and the size and structure of the firm. This direction provides perhaps the most natural extension for economic models of deterrence, particularly since the simplifying assumptions of utility maximization appear to be most relevant for studying business behavior as opposed to individual tax behavior. Given the relative lack of corporate compliance studies, Cowell's chapter would provide a major contribution to the compliance literature if it succeeds in fostering further studies of these important issues.

The section on strategic behavior reviews a few tax studies that have already begun to incorporate the strategic interaction between taxpayers and the tax collector into deterrence models. The models Cowell reviews do not appear to provide as a strong case for the relevance of this line of work as one might expect, but he suggests some interesting conclusions that caution against overly simple approaches to deterrence that ignore the agency's behavior. In particular, increasing enforcement stringency may induce strategic taxpayers to migrate to the "ghost" sector, or the underground economy, where their tax noncompliance may be even more costly and difficult to control.

The least developed but potentially most important new direction is provided in the section on rethinking taxpayer motivation. This section explores the possibility of including taxpayer motivations that have been found to be important in a wide range of compliance studies but that are traditionally excluded from economic analyses or treated as residual effects. Cowell briefly reviews four promising approaches: Utilize concepts like "framing" from behavioral decisionmaking studies to relax behaviorally questionable assumptions of full-information utility maximization; include the value of public goods in individual utility functions; expand
the time dimension of the ongoing relationship between taxpayer and tax collector; and explicitly model the endogenous interactions among taxpayers to reflect that “people may care about their own behavior relative to those of their peers.” Cowell goes on to note that “The maintenance of a culture of compliance is one example of the government or tax administration creating a ‘carrot’—a positive incentive for taxpayers to act in their broad social interest rather than in their narrow self-interest.”

This concern with incorporating a more relevant motivational framework for understanding taxpayer behavior is widely shared among compliance scholars. For example, a recent comprehensive review of tax compliance research concluded that “adding moral and social dynamics to models of tax compliance is as yet a largely undeveloped area of research. There seems to be little dispute [that] these factors are important in individual compliance decisions, but little is known or agreed upon about how best to include these effects in a theoretical or empirical analysis of tax compliance.”¹ My remaining comments expand on Cowell’s suggestions, in an attempt to incorporate concerns with justice into the utility-maximizing model, to make the model directly relevant to the current crisis of tax administration.

Deterrence or Assurance?

Although the extensions of utility-maximization models presented by Cowell can conceivably help tax administrators design and improve enforcement policies that enhance deterrence, the major changes in IRS procedures in the past decades are oriented more toward assurance than deterrence. The expansion of taxpayer services, taxpayer problem resolution offices, the taxpayer bill of rights, and the customer service orientation toward taxpayers are all intended to assure honest taxpayers that they will be treated fairly as long as they carry out their responsibilities as citizens. Combined with the IRS’s traditional enforcement activities intended to punish those who do not meet their tax obligations, the new activities are intended to assure that taxpayers receive their just deserts. The implicit assumption behind the growth of these new IRS activities is that the primary role of deterrence is to support assurance. Honesty is treated with respect, while dishonesty is punished. We need a model of compliance that recognizes both assurance and deterrence and that links both types of taxpayer motivations to tax policy and tax administration.

The emphasis on assurance is implicit in many of the views of practicing enforcement officials. Consider, for example, Chester Bowles’s perspective as director of the Connecticut Office of Administration Control during the Sec-

ond World War. He noted that the agency had little effect on 20 percent of people that would always obey the agency’s laws and a smaller group of “bad apples” that would never comply unless dragged into court. However, the large majority of people could be convinced to comply voluntarily, but only if the law was rigidly enforced enough of the bad apples to assure the average person that they were not “suckers” if they met their legal obligation to obey. Translating into the tax context, deterrence is most relevant for minimizing violations among the small group of hard-core evaders, whereas assurance is critical for maintaining voluntary compliance levels for the broadest group of taxpayers.

The model of compliance we need to develop is not really concerned with positive rewards (a lottery ticket as a reward for a clean audit), but rather with ensuring that taxpayers get what they deserve. “Just deserts” provides a better metaphor than the traditional image of a carrot as a positive incentive. To fully develop the model requires an extension into the literatures on clubs, contracts, collective action, and evolutionary psychology. I provide a brief overview here of the intuition behind the argument developed more fully elsewhere.

**Contractual Compliance**

The basic argument is that taxpayers obey the terms of an implicit tax contract as long as other taxpayers and the government meet their expected roles. The role of the general government is to produce worthwhile public goods, but the more specific role of the IRS is to ensure that other taxpayers cannot free-ride by shirking their tax obligations. This important assurance role requires the coercive powers of the deterrence model. But taxpayers also need to be assured that the coercive powers delegated to ensure the compliance of other taxpayers will not be misused to exploit honest taxpayers, whether from malevolence or organizational ineptitude. Administrative activities associated with customer service for taxpayers and the taxpayer bill of rights can be designed to fulfill this function.

How plausible is this model? Consider first the extreme case in which a predatory government attempts to extract the maximum possible tax from an understandably reluctant population. This is indeed the assumed model behind deterrence theory, since only the threat of punishment is expected to induce compliance. The problem with this model is that deterrence is expensive for both ruler and ruled, given the taxpayer’s incentive and ability to hide whatever is being taxed. Levy has argued that even in this extreme case, both rulers and the ruled can be better off if they can agree to reduce the deadweight loss involved in deterrence. If the king reins in the intrusive tax collector in return for

the barons’ “voluntary” contributions of taxes, both can be better off under this new contractual arrangement than they were under the repressive deterrence system. The king saves the cost of maintaining the coercive tax mechanism, while the barons save the costs of hiding their taxable assets from the intrusive tax collector.

Of course the new tax contract will be worthless unless both sides can provide some credible commitment toward fulfilling the contract. Levy argues that contingent compliance provides the critical foundation of this contractual system. If the king imposes additional taxes based on the now-visible assets of the barons, the barons can react to this breach of contract by withdrawing their voluntary contribution. The anticipated loss of revenue during these battle periods must be sufficient to cure the king of the temptation to cheat the barons. Similarly the king must maintain sufficient coercive powers to be able to punish barons if they fail to live up to their contractual obligation. In game theoretic terms, it is the contingent withdrawal of cooperation in repeated play of the game that provides the credible commitment to maintain compliance with the contract on both sides.

The predatory government perspective on contingent compliance, like the model of deterrence, captures only part of the picture of tax compliance for democratic governments, since democratic governments presumably provide desired public goods of value to taxpayers. To the extent that this is true, democratic taxation is closer to the theories of teams and clubs than to theories of predatory government. Both clubs and teams impose obligations on members in exchange for benefits derived from joint activities, just as democratic governments impose tax obligations in exchange for desired public goods. Thus teams, clubs, and government suppliers of public goods can be conceived of as implicit or explicit contracts that define the benefits and obligations of members or taxpayers.

The necessary condition for such a contract is that benefits exceed obligations for each member. But joint gains are insufficient without some means of assuring members that the contract will be enforced for all other parties, and the deadweight loss of costs required to enforce the contract is perhaps the greatest barrier to the broader provision of public goods. Thus the central problem in club and team theories is to devise efficient mechanisms for overcoming the problem of free riding that would otherwise prevent members from enjoying the benefits of joint production.

When members who fail to meet their obligation can be readily identified and excluded, as in some private clubs, there is little problem in obtaining the public goods. The problem arises when obligations are difficult to monitor and punish. An internal IRS could be set up to monitor and enforce the contract, but the cost of this deterrence approach is generally high and not necessary. For example, a simple strategy of shirking whenever team production falls below
some threshold can provide the necessary credible commitment to enforce a joint production contract without requiring the expense of monitoring individual behavior.\(^5\) As long as shirking has a sufficient probability of causing below-threshold production, and as long as the drop in joint production when everyone shirks is sufficient to overshadow the short-term gains of free riding, no team member would have an incentive to shirk his obligations.

Axelrod and others have demonstrated the power of reciprocity for maintaining cooperation, and have pointed out the importance of such contingent, reciprocal behavior in supporting a broad array of cooperative arrangements.\(^6\) As noted earlier, the baron's contingent compliance with the king's tax allows both king and baron to minimize the deadweight costs of enforcing their implicit tax contract. Miller applies this same argument to the relationship between managers and workers.\(^7\) He argues that corporate managers rely on reciprocity to maintain higher levels of productivity than could be achieved by more elaborate and costly enforcement mechanisms and incentive schemes. Managers eschew nit-picking monitoring and punishment of minor lapses, in return for expected efforts and flexibility on the part of the worker.

The advantages of cooperation have been so fundamental to human society since its early beginnings, according to Cosmedes and Tooby, that specialized human cognitive mechanisms have evolved to support the gains of contractual compliance.\(^8\) By developing cognitive mechanisms of trust, commitment, reciprocity, vengeance, and other behavioral strategies that can increase the credibility of contractual commitments, contemporary citizens are “better than rational” in their ability to resolve free-rider problems in ways that are beyond the reach of “rational fools.”\(^9\)

**Are Taxpayers Better than Rational?**

Taxpayers have generally been treated as rational fools because of the apparent intractability of the free-rider problem as applied to millions of taxpayers. For most analysts the rationality of contingent compliance is just too implausible, because taxpayers are unlikely to believe that their own behavior can have any impact whatsoever on other taxpayers. Without the IRS to provide a plausible assurance that other taxpayers will fulfill their obligations, any other concerns of the taxpayer are unlikely.

---

\(^5\) See, for example, Radner (1986.)

\(^6\) Axelrod (1984).

\(^7\) Miller (1992).

\(^8\) Cosmedes and Tooby (1994).

\(^9\) See also Frank (1988) for a delightful analysis of the instrumental utility of moral sentiments.
Scholz and Lubell on the other hand, argue the implausibility that citizens use cognitive mechanisms for tax obligations that are different from the mechanisms they use for the broad array of obligations from the many teams, clubs, and governing institutions they encounter in the rest of their lives.\(^{10}\) Think, for example, of the way individuals determine their obligation to keep implicit contracts with friends, teams, or clubs they belong to. Instead of storing every incident relevant to the obligation, people simplify the cognitive task by translating each incident or any relevant information into a series of overall attitudes and evaluations about the friend or club. When required to fulfill the relevant obligation, these attitudes determine the likelihood that the obligation will be met. The general evaluation of obligation for the particular friend or club provides an efficient decisionmaking process, leading to compliance when others are fulfilling their contractual obligations and noncompliance when they are not.

For tax compliance, this suggests a model in which taxpayers unconsciously process relevant incidents and information related to their income tax obligations into compliance-related beliefs and attitudes, just as they do with other laws. These attitudes subsequently lead to different probabilities of compliance in reporting taxes. For example, Scholz and Lubell interviewed taxpayers before and after they filed their first tax return affected by the 1986 Tax Reform Act (TRA). Taxpayers with large increases in tax due to the TRA systematically reduced their perceived obligation to pay taxes, suggesting that they are sensitive to the relative costs and benefits derived from the collective.\(^ {11}\) Furthermore, they found that self-reported tax compliance fell when obligation and trust in other taxpayers to pay their full share fell.\(^ {12}\)

While we know little about the actual cognitive mechanisms relevant for determining compliance with tax or other governmental obligations, Tom Tyler’s studies of compliance with police and court orders provide evidence of one particular mechanism that is directly related to the current crisis in tax administration.\(^ {13}\) Controlling for the outcome of their case, individuals subjected to police and court orders who feel that they have been treated with fairness and dignity are much more likely to comply than those who feel ill-treated by the system. Expectations about procedural justice provide one plausible mechanism by which individuals can evaluate the extent to which large, distant organizations like police, courts, and tax collectors have lived up to their contractual obligations.

Just as the medieval baron’s compliance may have been contingent on the expected behavior of the monarch, so also may the less-exalted taxpayer in con-

---

10. Scholz and Lubell (1998a,b).
temporary democracies respond contingently when IRS behavior falls short of socially determined expectations of procedural fairness. In contemporary society at least (and probably for the barons as well) the response is not a reasoned, conscious attempt to restrain the power of the tax collector. But the unconscious heuristic triggered by perceived breeches of procedural justice is part of a repertoire of cognitive and social mechanisms that provide the function of supporting contractual compliance by making compliance contingent on the behavior of others.

There is a good reason why procedural justice may be a critical part of citizenship responses, particularly for obligations imposed by large state agencies. Barzel argues that the consolidation of coercive enforcement powers in the state provides tremendous efficiency by enforcing standardized “contracts” that can be used by teams, clubs, and other organized joint production efforts in order to avoid the costs of creating their own coercive enforcement mechanism. However, a state enforcement agency with coercive power capable of ensuring compliance of even the most powerful individual is also capable of using that power to exploit members. Thus a critical problem for democratic governments is how to design effective constraints on the power of coercive enforcement agencies without destroying their ability to assure the credibility of contracts. Congressional oversight, judicial review of administrative procedures, the separation of tax enforcement from military and police agencies, and mandatory internal appeals processes are some of the institutional constraints on IRS authority. Taxpayer responses to procedural justice issues may provide a less familiar constraint that may be equally important in maintaining a tax system capable of supporting contractual compliance with tax obligations.

Implications for Tax Administration

If contractual compliance is relevant to tax compliance, we would expect that the level of compliance among taxpayers would be contingent on the perceived activities of the government and of other taxpayers. Compliance would decline if the tax collector became too intrusive or the government too exploitative. Thus an overzealous IRS would reduce compliance, and the reforms from taxpayer service to the taxpayer bill of rights would provide a means of redressing the balance and restoring the conditions of contingent compliance. On the other hand, if other taxpayers appear to be cheating on their taxes and the IRS appears to be too weak to catch and punish them, compliance would also fall. Critics of the customer-service orientation toward taxpayers argue that the past

decade of reforms has emasculated IRS enforcement capabilities to the point that it will soon be incapable of providing the needed assurance that other taxpayers will pay their taxes, resulting in a dramatic drop in compliance rates.

What is the appropriate balance between reforms favoring procedural justice and those that enhance enforcement effectiveness? The model of contractual compliance would suggest that the balance depends on which threat to assurance is causing the greatest concern among taxpayers. Unfortunately too little is known empirically about factors affecting contractual compliance to provide clear resolution of the current uncertainty between enhancing procedural justice and increasing enforcement effectiveness. The cautious note that Cowell emphasizes for his suggested extensions of compliance models applies even more fully for the contractual compliance model.

However, it is tempting to speculate about what a fully developed and empirically tested model of contractual compliance might offer in the way of advice to tax administrators. Use the “sticks” of deterrence where they are most appropriate—among subpopulations in which there is little support for contractual compliance. This ensures “just deserts” by providing maximal deterrence where contractual compliance has failed. But it also ensures that those willing to comply with the implicit tax contract are not discouraged from doing so by perceived injustices in their treatment by the IRS. Reduce compliance costs and burdens but in particular ensure that expectations about procedural justice are met as fully as possible whenever taxpayers are contacted. The appropriate balance of deterrence and procedural justice required for optimal assurance evolves over time with the nature of both the government and the governed and must be adjusted accordingly. Empirical and theoretical analyses of assurance and contractual compliance could help understand current challenges and provide a broader foundation for debate.

COMMENT BY
Douglas A. Shackelford

Chapter 8 reviews the prevailing economic model for noncompliance, details its weaknesses, and then discusses some possible improvements. It highlights the assumptions that underlie the model and explicitly details the propositions that fall out. Besides the theoretical model, the review also discusses the difficulties and weaknesses in empirical archival studies (for example, sample-selection bias) and includes a brief survey of tests conducted by experimentalists in economics and psychology. It is an excellent introduction for anyone trying to understand noncompliance.
The fundamental model, which the author terms TAG (taxpayer-as-gambler), treats the compliance decision as a classic trade-off between risk and taxes. The taxpayer opts either for a certain tax today or an uncertain tax tomorrow, after a possible government audit. Tomorrow’s tax may be lower, because the taxpayer intends to understate his taxable income. However, there exists a possibility that tomorrow the government will discover the understatement and not only force the full tax to be paid but add an additional penalty. I have found this model to be useful in structuring my thinking about compliance and, like many others, including several in this volume, I have used it in my own research.

The model, however, is not limited to compliance choices. You could imagine many settings where taxpayers face either a certain tax today or an uncertain (potentially lower) tax tomorrow. For example, consider a model where the taxpayer faces either a certain short-term capital gains tax today or a potentially lower long-term capital gains tax tomorrow. Although waiting until the property qualifies for the more favorable long-term treatment lowers the tax bill, it may result in a drop in after-tax profits, because there is a possibility that the value of the property will fall in the interim. In other words, the TAG model is a special case of a more general model that grapples with the coordination of taxes and risk.

As the author states, TAG assumes “rational individuals with stable preferences who, given economic opportunities and probabilities, maximize their expected utility.” It is amoral; it is nonstrategic; it ignores time. Like all theory, abstraction is necessary in order to construct a tractable, useful model. TAG ignores much of the economic, social, and philosophical richness that makes the compliance decision so interesting and so difficult to understand. The question is not whether TAG captures the “real world.” It clearly does not. The question is whether the more salient features of the compliance decision are included in the model.

The author identifies three “distinctly unsatisfactory” problems with the TAG model: its reliance on expected utility as the taxpayer motivation; its nonstrategic, atemporal nature; and its omission of the behavior of firms. I concur with the author that the model is both missing these features and would be improved if they could be added. Clearly compliance involves more taxpayer motives than simply expected utility. My observation is that some individuals take unusual pleasure in saving tax dollars, while others find legal tax avoidance vaguely inappropriate.

Likewise compliance decisions clearly are strategic. Both taxpayers and the government consider the multiperiod nature of compliance. The compliance model would be strengthened if it could be expanded to incorporate all economic agents. By not including these important aspects, the TAG model is clearly an incomplete, imperfect window on the world.
Besides these three arguably obvious limitations, the author adds that the nature of the product, the size and organizational structure of the firm, the role of reputation, and the degree of concentration of the industry also may affect compliance. Raising these more subtle considerations, which have been largely ignored in the literature, is a particularly useful contribution of the chapter.

Of course, if it were easy to incorporate these features into the noncompliance model, we would not be having this discussion. Thus apparently including these features comes at a high cost. So the real question is: What are the analytical costs of incorporating these benefits of realism in the model?

My disappointment with the chapter came at this point. It is unclear to me what to do with these problems or how to do it. If the purpose of the essay is simply to tell us that the TAG model is incomplete, that may be an important contribution. But I already knew that; I was looking for more guidance on where to go from here.

Since I do not know the way home either, I will join in the TAG bashing. In particular, I would like to see three improvements in the ultimate compliance model.

The Role of the Tax Preparer

Tax returns (personal or business) of any level of complexity involve a preparer. A complete model of noncompliance needs to consider the preparer’s role. One can argue that preparers improve compliance: They know the law better; they know the penalties (including special penalties on them) better; they know where to find answers to questions better. So one could argue that introducing a preparer lowers the level of noncompliance.

Alternatively preparers may increase noncompliance. Presumably they are paid with the proceeds from the tax savings they create. For example, corporate tax shelters are prima facie evidence that preparers increase noncompliance. Preparers know the legal ways to restructure transactions to achieve lower taxes; they know which options are likely to trigger audits and which are not. Consequently they can carefully select from a menu of avoidance options, choosing those that are most effective at lowering taxes and avoiding those that are least likely to be detected under an audit. The result is that the game is not between the government and the taxpayer. It is between the government and the taxpayer’s agent, who is an expert at playing the game.

The preparer’s role in compliance leads to classic principal-agent problems. For example, preparers face special preparer penalties if returns are found to be fraudulent. Thus, two penalties are considered in the production of the return—the taxpayer’s penalty and the tax preparer’s penalty. Also preparers face
both the cross-temporal issues raised in the chapter (choosing an avoidance option this year affects previous and future options), but also cross-client problems. Providing an avoidance option for one client affects the risk of audit for other clients using the same technique.

This relates to problems with property rights associated with tax plans. For example, the shelf life for corporate tax shelters is short. With each application of the shelter, the probability of its continuing usefulness diminishes. Presumably preparers ration noncompliance options across taxpayers in a manner that maximizes the preparers’ profit.

One promising option for improving TAG is to incorporate findings from the accounting literature concerning the role of tax preparers. Academic accountants have a longstanding interest in the role of tax preparers (who, of course, are often practicing accountants). Roberts and Cloyd and Spilker provide examples of the research in this area.¹ Not surprisingly these studies find the taxpayer-preparer-government triangle to be complex.

The Political Environment

The second dimension that I would like to see incorporated into the ultimate compliance model is the political environment. I have no idea how to do this, but it is a major factor in understanding the seemingly increasing level of noncompliance.

What is the genesis of corporate (and increasingly personal) tax shelters? Any discussion of noncompliance requires an inquiry into these shelters. Their size alone demands attention. The New York Times quotes a tax attorney at Sullivan and Cromwell who states that “the government needs to devote ten times as many resources as it does now if it wants to tax capital effectively.”² Where did this monster come from? If we want to understand corporate tax shelters, we must understand the political environment in which they have arisen. Many will point to changes in federal government over the past decade, such as the IRS-bashing of 1997 and 1998, for a political explanation for the emergence of these shelters (and I would not disagree). However, I think it is useful to move back a few more years to some unintended consequences of deregulation.

One group that many label as a primary source of the noncompliance problem is the Big (or Final) Four accounting firms. The financial statements of all

¹. Roberts (1998); Cloyd and Spilker (1999).
publicly traded firms must be audited annually. The size and scope of large multinationals limit their potential auditors to the largest multinational accounting firms (of which only the Final Four remain). Let me briefly review the evolution of these firms.

In the late 1970s the auditing industry, as well as several other professions, underwent what I will call deregulation. Until then, the code of professional conduct for the American Institute of Certified Public Accountants had forbid all forms of advertising, including soliciting business from another firm’s client without first asking their auditor’s permission. Over the quarter century since the elimination of these barriers to entry, the profit margins on audits have shrunk considerably. (Bankman’s chapter in this volume terms these returns “unspectacular.”) The result has been less auditing and movement into higher-margin businesses, such as consulting. It is a long story, but a line can be drawn from deregulation to lower fees to Enron, WorldCom, and other debacles that have left the auditing industry in shambles.

What does this have to do with tax noncompliance? Conducting an audit requires the collection of much of the same information that is required to complete an income tax return. Over time the accounting firms began to provide tax advice and complete tax filings for their audit clients. After deregulation, firms began to compete for audit and tax assignments, driving down the fees for both services. Firms also began to compete for profitable tax consulting assignments for nonaudit clients. This competition led to a further reduction in the profit margins on traditional tax work. Thus, in the same way that auditors began to look for higher-margin activities to replace the diminishing returns from auditing, tax advisers began to look for more profitable tax plans.

Twenty-five years later, the most profitable tax product for multinational corporations is a loose collection of tax plans known as corporate tax shelters. These shelters generally meet the letter but not the spirit of the law. I would term them legal noncompliance. To some extent noncompliance attributable to the large accounting firms is an unintended consequence of deregulation from a quarter century ago. This should not be construed as regret that trade restraints were lifted but rather recognition that the lifting indirectly changed tax compliance. I present the evolution of the tax industry as evidence that a complete model of noncompliance needs a thorough dose of political economics.

Public Disclosure of Tax Liability

Finally, to fully understand compliance for publicly traded corporations, we need to consider the information about taxes that the public receives. Consider Enron or WorldCom. Did they pay taxes on the profits they reported but never earned? If so, then we have noncompliance that results in tax overpayment. Erickson,
Hanlon, and Maydew attempt to address this question.3 They show that companies actually paid taxes on allegedly fraudulent earnings. That is, the profits they reported, but never actually earned, were reported as taxable income. Their paper builds on a long line of studies showing that publicly traded companies often forgo tax avoidance opportunities that result in reduced earnings, because they value the accounting earnings (and their impact on share price) more than the cash outlay associated with the taxes.4 So, if we wish to understand tax compliance for publicly traded companies, we need to consider the information about profits and taxes paid that these companies must provide investors and recognize that the public nature of this information may lead to increased tax revenue.

In short, this chapter is an excellent primer for understanding compliance issues. Both theorists and empiricists can benefit from its insights and its guidance for future compliance research. The challenge is to develop models that better capture the richness of the compliance decision.

References


4. See review in Shackelford and Shevlin (2001)


CARROTS AND STICKS IN ENFORCEMENT


