

Knowledge of the Tax Law*

David Weisbach[†]

Preliminary. Subject to Revision.

February 7, 2011

Abstract

This paper considers whether knowledge of the tax law is socially desirable. Unlike other laws, which most often attempt to channel behavior, revenue raising taxes attempt to avoid changing behavior, so it is not obvious whether or when knowledge of the tax law is desirable. I argue that whether knowledge of the tax law is desirable depends on three factors: expectations about the tax in the absence of knowledge, the type of tax, and the quality of the tax. I then apply this to various policy decisions where knowledge of the tax law is a key variable, including the regulation of tax advisors, hidden taxes (such phase-outs of exemptions), and whether individual have the socially optimal incentive to seek knowledge of the tax law.

1 Introduction

In most cases, knowledge of the law is desirable. We want people to know how to dispose of hazardous waste, to enter into binding contracts, and to demarcate their property. People should know the extent of their free speech rights and that assault is criminal. Tax law, however, may be different, at least with respect to knowledge of the law prior to taking action. Often tax

*This work was supported by __.

[†]Walter J. Blum Professor, The Univesity of Chicago Law School. I thank Brian Galle, Jacob Goldin, Saul Levmore, Yair Listokin, Mark Phillips, and participants at the Loyola Law School Tax Colloquium for conversations and comments. email: d-weisbach@uchicago.edu

law tries to achieve its goals – raise revenue, redistribute, and so on – while minimizing changes to behavior, and behavioral changes may be minimized if people do not know the details of the law. For the tax law, ignorance may be bliss, if not privately, at least socially.

Understanding whether and when knowledge of the tax law is desirable is central to a number of important policies. For example, tax systems can be designed to be easier or harder for people to understand and incorporate into their behavior. A number of commentators have suggested that it might be desirable to purposefully choose tax systems that are difficult to understand on the theory that such systems will have lower efficiency or related costs.¹ Thus, a VAT builds the tax into the price of a good so that buyers will automatically take account of the tax in their purchasing decisions.² A retail sales tax, which, if taxpayers have full knowledge of the law can be made economically equivalent to a VAT, may be less transparent because it is imposed at the cash register; the prices of goods are typically displayed without the tax. The cost of determining the after-tax price under a retail sales tax is higher than with a VAT which means that taxpayers are less likely to fully understand the impact of a retail sales tax than a VAT. If knowledge of the tax law is undesirable, a retail sales tax, all else held equal, becomes relatively more desirable. Alternatively, hiding the marginal rate structure within an income tax, such as through non-transparent phase-ins and phase-outs, may be desirable.

Similarly, in the last 20 years there has been an explosion of rules regulating tax lawyers, increasing the cost of tax advice and reducing its value.³ The new rules regulate the content of tax advice, limit the allowable fee structures, reduce the ability to keep tax advice confidential, and limit the value of tax advice by reducing the cases where taxpayers can rely on it to avoid penalties. Understanding whether knowledge of the tax law is socially

¹Jeffrey B. Liebman and Richard J. Zechhauser, *Schmeduling*, Unpublished manuscript available at <http://www.hks.harvard.edu/jeffreyliebman/schmeduling.pdf> (October 2004) (arguing that the tax system can be more efficient if people systematically underestimate marginal rates); Deborah Schenk, *Exploiting the Salience Bias in Designing Taxes*, NYU Law and Economics Research Paper No. 10-37 (August 18, 2010) (suggesting the use of low salience provisions in situations where the legislature is inappropriately limited in its ability to raise marginal rates).

²For a discussion of the mechanics of VAT and how they relate to retail sales taxes, see David Weisbach, *Ironing Out the Flat Tax*, 52 *Stanford L. Rev.* 599 (2000).

³See Brian Gale and David Weisbach, *The Regulation of Tax Advisors*, forthcoming, *Tax Notes*, March 2011.

desirable is central to evaluating these rules. If knowledge of the tax law is undesirable, we may want to limit access to tax lawyers or otherwise raise the price of tax advice while, if knowledge of the tax law is desirable, these conclusions may change.

Yet another important tax policy decision affected by this question is the extent to which the tax law can or should be uncertain. Many authors, including me, have advocated for uncertainty in the form of broad-based standards designed to prevent or reduce tax shelters.⁴ In 2010, in part in response to these arguments, Congress enacted a statutory economic substance doctrine, which overlays a broad anti-abuse standard on top of the existing detailed tax rules.⁵ A criticism of the economic substance doctrine is that it makes it too hard for people to know what the law is.⁶ If knowledge of the tax law is undesirable, however, this criticism is muted or possibly even completely flipped.

The list goes on. Tax patents, which are a subset of business method patents, provide intellectual property protection for tax innovations.⁷ We can view tax innovations as new knowledge of the tax law. Therefore, whether tax patents are desirable may then turn on whether knowledge of the tax law is desirable. And there is a large literature on tax transitions, most of it arguing that tax changes should be retroactive.⁸ A concern with retroactive changes, like with broad-based standards, is that they introduce uncertainty. If the law can change retroactively, people cannot easily determine what the law is. An understanding of whether this uncertainty imposes costs or is actually desirable, can inform the debates on retroactivity.

Policy decisions in each of these areas, and others, depend on detailed considerations particular to the area. For example, some argue that uncertain tax laws lead to excessive ex post litigation costs.⁹ If so, this may outweigh considerations of whether knowledge of the law ex ante when people are

⁴David Weisbach, *Formalism in the Tax Law*, 66 Chicago L. Rev. 860 (1999).

⁵See Internal Revenue Code section 7701(o) (2010).

⁶See, e.g., Kenneth W. Gideon, *Assessing the Income Tax: Transparency, Simplicity, Fairness*, 81 Tax Notes 999 (November 23, 1998).

⁷See David Weisbach, *Disrupting the Market for Tax Planning*, 26 Virginia Tax Review 971, 974 (2007).

⁸See, e.g., Michael Graetz, *Legal transitions: The case of retroactivity in income tax revision*, 126 Penn L. Rev. 47 (1977); Louis Kaplow, *An economic analysis of legal transitions*, 99 Harv. L. Rev. 509 (1986), Daniel Shaviro, *When Rules Change* (2000).

⁹Kenneth W. Gideon, *Assessing the Income Tax: Transparency, Simplicity, Fairness*, 81 Tax Notes 999 (November 23, 1998).

deciding what actions to take is desirable. The same considerations, however, may not apply to strict regulation of tax advisors, where other issues may instead be important. Nevertheless, whether knowledge of the tax law is desirable is a central variable to thinking about these questions.

In this paper, I ask whether knowledge of the tax law prior to taking action is desirable. Is tax system is more efficient when taxpayers know the law prior to acting or when they do not? As a secondary matter, I also consider whether people have socially optimal incentives to seek knowledge of the tax law prior to taking action. If knowledge of the tax law is socially undesirable, then it is likely that individuals have excessive incentives to seek such knowledge, but their incentives may be socially excessive even if knowledge of the tax law is desirable.

A central limitation to this framing is that it is limited to whether knowledge of the tax law is desirable prior to individuals taking action. Knowledge of the tax law for purposes of return preparation (after taking action) may raise different issues.¹⁰ Ex ante and ex post knowledge, however, will be hard to separate in practice because ex post knowledge on one year serves as ex ante knowledge for the next. Therefore, in determining actual policies, both have to be taken into account. Nevertheless because ex ante and ex post knowledge of the tax law raise conceptually distinct issues, it is worth separating the two for purposes of analysis.¹¹

¹⁰See, e.g., Steven Klepper, Mark Mazur, and Daniel Nagin, *Expert Intermediaries and Legal Compliance: The Case of Tax Preparers*, 34 J. L. & Econ. 205 (1991) (modeling tax preparers and taxpayers and concluding that increased penalties help decrease noncompliance on legally ambiguous items but also raise costs to taxpayers, leading to less use of preparers); Jennifer F. Reinganum and Louis L. Wilde, *Equilibrium Enforcement and Compliance in the Presence of Tax Practitioners*, 7 J. L. Econ. & Org. 163 (1991) (constructing economic model of taxpayers, practitioners, and agents to explore linkages between use of practitioners and noncompliance by taxpayers); Steven Klepper and Daniel Nagin, *The Role of Tax Practitioners in Tax Compliance*, 22 Pol'y Sci 167 (1989) (explaining how tax preparers serve two roles: (1) increasing tax compliance in unambiguous areas; and (2) reducing compliance ambiguous areas); Suzanne Scotchmer, *The Effect of Tax Advisors on Tax Compliance*, reprinted in *Taxpayer Compliance: Social Science Perspectives*, Volume 2 182 (Jeffrey A. Roth & John T. Scholz eds., 1989) (asserting that resolving taxpayer uncertainty increases utility by reducing risk but also may reduce net revenue).

¹¹Saul Levmore proposed the following conceit to separate the two: the tax law is written down at the beginning of each year but is sealed in an envelope. At the end of the year but before returns are filed, the envelope is opened. Taxpayers would have full information about the law for purposes of filing returns and no or at least less knowledge of the law

I will argue here that whether knowledge of the tax law is desirable depends on three factors: (i) beliefs about the tax law in the absence of knowledge; (ii) the quality of the tax law (i.e., how well designed is it?); and (iii) the type of tax. Beliefs matter because in the absence of knowledge, they determine behavior. To determine whether knowledge of the tax law causes more or less efficient behavior, we have to compare behavior with and without knowledge. Behavior without knowledge of the tax law will depend on expectations and beliefs, which means that they are central to the analysis. The quality of the law matters, because if the law is optimally designed, an understanding of the law is far more likely to lead to efficient behavior than if the law is badly designed. Finally, the type of tax matters. The simplest example of why the type of tax matters is the difference between a Pigouvian tax and a tax designed to raise revenue. Pigouvian taxes internalize externalities, making them similar to tort law. It is likely desirable that people know about Pigouvian taxes, while the same intuitions may not carryover to revenue-raising taxes. Similar considerations apply with different types of revenue raising taxes, such as linear compared to nonlinear taxes.

Unfortunately, these factors are often hard to observe. Moreover, subtle changes in these factors can change the conclusions, so accurate observation is necessary. The result is that it will hard to know in any given case whether knowledge of the tax law is good or bad.

This suggests that we should limit policies that depend on whether knowledge of the tax law is desirable to circumstances where we are comfortable that we understand how the three factors interact. For example, Raj Chetty and Emmanuel Saez conduct an experiment in which they educate low-income taxpayers about the earned income tax credit.¹² As will be discussed below, this might be a situation where we are comfortable that knowledge of the tax is desirable. Similarly, we might regulate tax advisors more heavily in areas where we believe that knowledge of the tax law is not desirable.

In other cases, where the empirical estimates or interactions of the relevant factors are unclear, perhaps we should not base policies on whether knowledge of the tax is desirable. Therefore, we should not, as a general matter, try to manipulate salience. Similarly, if uncertainty in the tax law has good effects, perhaps we should not worry about whether uncertainty

for purposes of choosing their actions.

¹²Raj Chetty, and Emmanuel Saez, *Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients*, NBER Working Paper 14836 (2009).

makes it more difficult to know what the tax law is.

I am not aware of any papers that directly analyze whether knowledge of the tax law prior to acting is desirable. Nevertheless, there are a number of papers which are closely related, and I will use some of their results. The most closely related work is the recent work on salience. A number of scholars have presented evidence that certain types of taxes are not salient in the sense that taxpayers, even if they know about a tax when asked, often fail to take it into account when making decisions.¹³ Most of this work is empirical, attempting to measure which taxes are salient and which are not. Several of the papers, however, also examine the welfare effects of salience.¹⁴ Salience is similar to but not exactly the same as knowledge – people could know the law but it could still lack salience because of inattention. Nevertheless, it is sufficiently similar that the results of these papers will be central to the analysis here.

Another related literature is on accuracy in the tax law. Because taxpayers may have full knowledge of an inaccurate tax, the issue is not precisely the same. Nevertheless, we can think about inaccuracy similarly to uncertainty about the content of the law. The reason is that inaccuracy creates dispersions in outcomes. The effect, from an ex ante perspective, is similar to uncertainty. The most relevant papers here are by Louis Kaplow¹⁵ and by

¹³Papers include Raj Chetty, Adam Looney, and Kory Kroft, *Salience and Taxation: Theory and Evidence*, 99 *American Economic Review* 1145 (2009) (retail sales taxes and alcohol taxes not salient); Amy Finkelstein, *E-Z Tax: Tax Salience and Tax Rates*, 124 *Quarterly Journal of Economics* 969 (2009) (automatically collected highway tolls are less salient than normal tollbooths); Charles Bartolome, *Which Tax Rates Do People Use: Average or Marginal?*, 51 *Journal of Public Economics* 79 (1995) (individuals mistake average rates for marginal rates); Jeffrey B. Liebman and Richard J. Zeckhauser, *Scheduling*, Unpublished manuscript available at <http://www.hks.harvard.edu/jeffreyliebman/scheduling.pdf> (October 2004) (same); Tomer Blumkin, Bradley Ruffle, and Yosef Ganun, *Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment with Real Goods*, CESifo Working Paper No. 2194 (2008) (economically equivalent income and consumption taxes perceived as different); Jonathan Baron and Edward J. McCaffery, *The Humpty Dumpty Blues: Disaggregation Bias in the Evaluation of Tax Systems*, 91 *Organizational Behavior and Human Decision Processes* 230 (2003) (individuals underestimate the burden of disaggregated taxes); Rupert Sausgruber and Jean-Robert Tyran, *Testing the Mill Hypothesis of Fiscal Illusion*, 122 *Public Choice* 39 (2005) (buyers underestimate the burden of indirect taxes compared to a direct tax).

¹⁴In particular, Chetty, Looney and Kroft, *Salience and Taxation*, and Leibman and Zeckhauser, *Scheduling*, have welfare analyses that are used in sections 1 and 2 below.

¹⁵Louis Kaplow, *Accuracy, Complexity and the Income Tax*, 14 *Journal of Law, Eco-*

Joel Slemrod and Christian Traxler.¹⁶

Finally, there is a literature on uncertainty in the tax law, which is closely related to knowledge of the law.¹⁷ I have argued that the effects of uncertainty in the tax law are ambiguous.¹⁸ Moreover, there may be benefits to broad-based standards in that they allow the content of the law to be specified after individuals act rather than before.¹⁹ Others, however, have argued that uncertainty is damaging to the tax law.²⁰ There is a separate and somewhat distinct literature arguing that randomness is actually a good thing.²¹

Section 1 develops the arguments on whether knowledge of the tax law is socially desirable through a series of relatively simple examples. Section 2 considers the incentives to acquire knowledge of the tax law. Section 3 circles back and discusses the assumptions and basic economics used in the examples in a somewhat more formal fashion. Section 4 concludes.

nomics and Organization 61 (1998).

¹⁶Joel Slemrod and Christian Traxler, *Optimal observability in a linear income tax*, 108 *Economics Letters* 105 (2010).

¹⁷It is also worth noting that there is a literature focusing on knowledge of non-tax law, usually using torts as the motivating example. This literature is relevant here to the extent taxes are substitutes for tort or other laws. For example, Pigouvian taxes, which are taxes on harm causing activities, can be thought of as substitutes for tort liability. For some of the relevant literature, see, e.g., Kaplow and Shavell, *Private versus Socially Optimal Provision of Ex Ante Legal Advice*, 8 *J. Law, Economics, and Organization* 306-320 (1992), Richard Craswell and John Calfee, *Deterrence and Uncertain Legal Standards*, 2 *Journal of Law, Economics and Organization* (1986); Louis Kaplow, *Optimal Deterrence, Uninformed Individuals, and Acquiring Information about Whether Acts are Subject to Sanctions*, 6 *Journal of Law, Economics, and Organization* 93 (1990); Steven Shavell, *Legal Advice about Contemplated Acts: The Decision to Obtain Advice, Its Social Desirability, and Protection of Confidentiality*, 17 *Journal of Legal Studies* 123 (1988).

¹⁸David Weisbach, *Ten Truths About Tax Shelters*, 55 *Tax L. Rev.* 215, 247 (2002).

¹⁹David Weisbach, *Formalism in the Tax Law*, 66 *Chicago L. Rev.* 860 (1999).

²⁰Kenneth W. Gideon, *Assessing the Income Tax: Transparency, Simplicity, Fairness*, 81 *Tax Notes* 999 (November 23, 1998).

²¹See, e.g., L. Weiss, *The desirability of cheating incentive and randomness in the optimal income tax*, 18 *Journal of Political Economy* 1343 (1976), Joseph Stiglitz, *Utilitarianism and horizontal equity: the case of random taxation*, 18 *Journal of Public Economics* 1 (1982). Other authors argue that the Weiss and Stiglitz arguments are not convincing. See, e.g., Martin Hellwig, *The undesirability of randomized income taxation under decreasing risk aversion*, 91 *Journal of Public Economics* 791 (2007).

2 Simple Examples

In this section, I develop the intuitions regarding knowledge of the tax law through a series of simple examples. More detail on the arguments and assumptions that are made here can be found in section 2. The goal here is to keep the discussion informal and intuitive.

The examples show that there are three different factors that together determine whether knowledge of the tax law improves the efficiency of the tax system. Each of these factors can be present with any combination of the other factors. This means that there are a large number of different possibilities. To keep the discussion manageable, I only list seven.²² In each case, we want to consider how taxes distort behavior when individuals know all of the details of the tax and when they do not.

Example 1 *Optimal linear tax, unbiased expectations*

Start with an extremely simple world where there are only two goods, automobiles and food, and a single consumer who wants to have equal amounts of each. The tax is a sales tax on automobiles at 10% and on food at 2%.

Assume, for now, that these tax rates are optimal. In particular, assume it is desirable to tax automobiles at a higher rate than food. There could be many reasons for this. A familiar one is we may want to set the tax rates so that the change in the percentage demand for each good because of the tax is the same. It might take a higher rate on automobiles than on food to achieve this effect. (The example is just made up – it could equally be the other way around.) While there are a number of alternative reasons for imposing differential rates on different commodities, use this one as a working example.

If the consumer knows the tax rates, the consumer will adjust his behavior as desired. So, if achieving an equal percentage decline demand for each good is the reason for the differentiation, the consumer will reduce his demand for each good accordingly. If the reason for these particular tax rates is different, the desired goal will also be achieved simply by assumption that tax rates are set optimally.

If the consumer does not know about the tax, we have to specify what he assumes about it. Just because he does not have full knowledge of the

²²Depending on how one counts different situations, there are at least 12 possibilities and possibly many more.

specifics of the tax does not mean that he assumes it to be zero. Instead, he could assume the rates are higher or lower than they actually are, or on average about right.

A natural case to start with is the case where the consumer's expectations about the tax are unbiased. Suppose that the consumer has \$112 in his pocket and buys equal amounts of each good. He can purchase \$50 pre-tax of each item. If he believes that rates to be any combination that averages to 6%, then the tax he expects to pay and the tax he actually pays are the same. When he gets to the cash register, his expectations about his taxes are met, and he will have budgeted enough to pay the tax. For simplicity, and it will turn out to be without loss of generality, assume that the consumer believes both rates to be 6%.

Is the tax more or less efficient when the consumer knows the rates than when he has unbiased expectations about the rates, which under our assumptions means he assumes both rates are 6%? It is more efficient when the consumer knows the rates. This conclusion is built into the assumptions of the example – that the tax rate differentials were desirable. This means that behavior is more efficient when automobiles are taxed at 10% and food at 2% than when both are taxed at 6%. If the consumer behaves as if both were taxed at 6%, we lose those benefits.

For example, if at 10% and 2%, the percent demand for each good declines by the same amount, at a flat 6%, they would not. The demand for food would go down by too much and the demand for automobiles by not enough. We could improve the efficiency of the tax by increasing the (perceived) tax on automobiles and reduce it on food.

A slightly more technical version of this argument is that the efficiency losses from taxation go up with the square of the tax rate. (The reasons for this are discussed in Part __ below.) If the consumer believes that the tax rate on food is higher than it really is, there will be additional efficiency losses. If the consumer believes that the tax on automobiles is lower than it really is, there are efficiency gains. But the losses are larger than the gains because of the nonlinearity of the efficiency losses.

We can now see why the assumption that the consumer believed that both rates were 6% is without loss of generality. Any combination of rates that average to 6%, other than the optimal rates, 10% and 2%, will be less efficient than the optimal rates. This holds even if the consumer believes the tax rate on food is lower than it is (instead of higher) and the rate on automobiles is higher, such as if the consumer believes that the tax rate on

food is zero and the tax rate on automobiles is 12%.

Therefore, we can conclude in this initial case, that knowledge of the tax law makes the tax system work better. It reduces the efficiency losses from taxation.

Note however, that the unbiased expectations assumption is difficult to maintain when the taxpayer has incorrect views about specific tax rates. In the example, if the taxpayer believing that the tax rates are all 6% adjusts his behavior to spend more than half of his budget on automobiles, the average tax rate will not be 6%. The only way expectations can be unbiased is, given the taxpayer's beliefs about tax rates, after-tax behavior confirms those beliefs. In the numbers used, it has to be the case that given a belief that tax rates are the unweighted average of the actual rates, the taxpayer happens to spend equal amounts on both goods. This is not impossible. Given a particular spending pattern, a taxpayer may infer an average tax rate based on the taxes he pays. But if the taxpayer subsequently changes behavior, expectations would no longer be unbiased.

Example 2 *Optimal graduated tax, unbiased expectations*

Consider next a tax that is graduated instead of linear. For example, consider a tax on labor income with increasing (or decreasing) marginal rates. If marginal rates are increasing, a taxpayer's average rate will be below his marginal rate. For example, consider a simple two-bracket tax which imposes a zero rate up to \$30,000 of income and a 50% rate for income above that amount. If a taxpayer earns \$40,000, he owes \$5,000 in tax (half of his income above \$30,000). His average tax rate is 12.5% while his marginal rate is 50%.

If the taxpayer assumes his marginal rate is his average rate, he has unbiased expectations in the sense discussed above. He can meet his budget and does not learn that he has miscalculated when it comes time to pay the tax. He is still, however, assuming that the marginal rate is lower than it is. This means that labor decisions will be less distorted than they would be if he had full knowledge of the tax law.

Knowledge of the law in this case, therefore, appears to reduce efficiency, even though the law is set optimally and expectations about taxes are unbiased (in the sense that people actually pay the amount that they expect to pay). The reason is that marginal tax rates matter for efficiency. Someone deciding whether to work more cares about the tax on the marginal dollar. With increasing marginal rates, unbiased expectations means that people underestimate marginal rates.

Jeffrey Liebman and Richard Zeckhauser have argued that people in fact have these beliefs about the income tax, something they call “schmeduling” (for schmeared schedules).²³ They then calculate the benefit in terms of lower efficiency cost and conclude that the deadweight loss is less than half of what it would be if people knew their true marginal rates. The same is true for marginal deadweight loss, which more than doubles if we assume that people know their true marginal rates instead of schmeduling. Therefore, in this example, the loss from better knowledge of the tax law would be substantial.

Similarly, Raj Chetty and Emmanuel Saez conduct an experiment where they teach people (or more accurately have tax return preparers teach people) about the rate schedule for the earned income credit.²⁴ They assert that people know their average rate but not their marginal rate (p. 2). The earned income credit has declining marginal rates in the relevant range because the reimbursement rate goes up with income.²⁵ This means that deadweight loss should go down when people know their true marginal rates (which they over-estimate if they assume their average rate is their marginal rate). Chetty and Saez show substantial effects on work effort as a result of the instruction. Knowledge of the tax law is beneficial in this case. As a result, Chetty and Saez (p. 34) suggest the possibility of making the phase-in range of the earned income credit (where marginal rates decline) highly salient and the phase-out range (where marginal rates increase) less salient, in an attempt to make knowledge of the law cheap where knowledge is desirable and more expensive where knowledge of the law is not desirable.

Example 3 *Two taxpayers, graduated rates*

Suppose now that there is more than one taxpayer and the tax is a non-linear tax on labor income. Some taxpayers may be informed about the rate structure while others are uninformed. This may occur across income groups but may also occur for individuals with the same wage income or the same

²³Jeffrey B. Liebman and Richard J. Zeckhauser, *Schmeduling*, Unpublished manuscript available at <http://www.lks.harvard.edu/jeffreyliebman/schmeduling.pdf> (October 2004). See also Charles Bartolome, *Which Tax Rate Do People Use: Average or Marginal?*, 56 *Journal of Public Economics* 79 (1995) (experimental evidence that people use average rates).

²⁴Raj Chetty and Emmanuel Saez, *Teaching the Tax Code, Earnings Responses to an Experiment with EITC Recipients*, NBER Working Paper No. 14836 (April 2009).

²⁵See IRC section 32, and Chetty and Saez, Figure 1.

earning ability because some individuals may have a particular taste for (or against) learning the tax law.

As will be discussed below, if knowledge varies across income groups, it may have distributive effects in the normal sense of distribution from rich to poor. In this case, we can adjust the tax brackets to offset this effect. For example, suppose that the rich have better knowledge of the tax law than the poor, resulting in relatively lower payments for the rich given the nominal tax rates. We can always increase the nominal tax rates on the rich and lower the rates on the poor to offset this effect.

If knowledge varies within an income or ability class, we cannot easily adjust the rate structure to offset the distributive effects. To illustrate, consider two taxpayers with the same ability to earn income, one uninformed about marginal rates but unbiased (so he believes his marginal equals his average rate) and one informed. Suppose that the tax is on labor income with increasing rates.

The informed individual will reduce his labor income more than the uninformed individual because he understands his true marginal rate which is above his average rate. The informed individual is better off. He better optimizes, so his trade-off between labor and leisure is superior. In addition, if there is a fixed budget constraint, overall rates can be lower because the uninformed individual works too much. The informed individual gets the benefit, effectively transferring dollars from the uninformed individual to the informed individual.

We cannot easily evaluate these sorts of distributive effects. If we use a simple social welfare function that values redistribution based on the marginal utility of income, we should want to redistribute in the other direction. The uninformed individual will have lower utility because of his failure to optimize. As a result, his marginal utility of income will be higher. We should want to redistribute from the informed to the uninformed individual rather than the other way around. Nevertheless, an understanding of the reasons some individuals have information about the tax law and others do not may change the social evaluation; the uninformed individual may be able to spend his leisure time in better ways than studying the tax law. That is, we have to understand the source of the individual variation to understand its welfare consequences. Therefore, it is not straightforward to make welfare comparisons with this type of redistribution.

Example 4 *Pigouvian tax*

A third type of tax is a Pigouvian tax, which is a tax on a harm-causing activity such as pollution. A tax equal to the marginal harm from an activity causes actors to internalize the harm. We can think of Pigouvian taxes as alternatives to tort liability.²⁶ Therefore, the normal considerations for when knowledge of the law is desirable should apply rather than the considerations examined here for other types of taxes. In general, as for tort law, knowledge of the law is desirable.

Example 5 *Optimal linear tax, biased expectations*

A second way to vary the example is to drop the assumption that expectations are unbiased. People who lack knowledge of the tax law may systematically under or over-estimate the applicable rates.

To see the effect of biased expectations, go back to the simple sales tax on food and automobiles, at 2% and 10% respectively, but suppose that the consumer mistakenly assumes that both are taxed at 2%. There are two key effects and the net result – whether knowledge of the law increases or reduces efficiency – will depend on the relative magnitudes of these effects. First, because the consumer thinks the tax rate is lower than it is, he will choose a bundle of goods that is closer to the one he would have picked with prices undistorted by taxes. In our example, the perceived tax rate on food is the same as the actual and optimal rate. Therefore, efficiency losses from the tax on food are the same as in the full-knowledge case because the individual correctly perceives this rate. The perceived rate on automobiles, however, is lower than the actual rate, so the choice of how much to spend on transportation is less distorted.

This effect – the substitution effect – is somewhat counter-intuitive in these circumstances. It is, by assumption, optimal to tax automobiles at a higher rate than food. If both were actually taxed at 2%, we could improve the efficiency of the tax by lowering the rate on food and increasing the rate on automobiles while holding tax receipts constant. This thought experiment, however, does not apply in the present context because we are considering only perceived rates, not actual payments. Because we are allowing downward-biased perceptions, we can lower the perceived rate on

²⁶See Steven Shavell, *The Corrective Tax versus Liability As Solutions to the Problem of Harmful Externalities* (Harvard John M. Olin Center for Law, Economics and Business, Discussion Paper No. 672, 2010)

automobiles without implicitly increasing the rate on food.²⁷

The second effect is that the individual will, unfortunately, find that his budget does not balance. He owes more in taxes than he expected, and therefore, is poorer than he thought. Given that he has overall less income, he might not have chosen the same amounts of each good as he would have chosen had he known. For example, the relative consumption of one of the goods, say automobiles, might go up with income. Because the individual assumed he had more income than he really did, he might have spent too much on automobiles. Had he known his true budget, he would have chosen differently. This effect – the income effect – reduces the efficiency of the tax.

We cannot say, a priori, whether the net of these effects is more efficient than if the taxpayer knew the correct rates. It will depend on the taxpayer's beliefs in the absence of knowledge and the relative size of the two effects given those beliefs. If the taxpayer over-estimates tax rates, the two work in the same direction and knowledge of the law is desirable. If the taxpayer underestimates tax rates, the two work in the opposite direction and the net effect could go either way.

A third consideration is that the taxpayer will likely adjust his beliefs when he finds out he has incorrectly computed his taxes. Each time the taxpayer pays taxes, he is exposed to information about whether his beliefs are correct. If his beliefs are biased, he will find out. If the taxpayer eventually moves toward unbiased beliefs, the net effect of lack of knowledge, whether positive or negative, will be limited.

The analysis will be the same with graduated taxes, such as a graduated tax on labor income. The substitution effect means that the individual will choose how much to work based on the perceived rate, which could be higher or lower than the marginal rate. There will also be an income effect because the individual will have assumed he has a different amount of money than he actually has, so his decisions will be distorted.²⁸

Example 6 *imperfect linear tax, unbiased expectations*

The final way to vary the example is to assume that the tax is imperfect or outright bad. This might be because administrative or compliance costs force

²⁷Note that there could be cross-effects, in which the tax on automobiles affects purchases of food. The discussion in the text assumes that these are zero or small.

²⁸For Pigouvian taxes, knowledge of the law will continue to be good – if a taxpayer under or over-estimates the tax, he will not properly internalize the harm he causes.

us to use simple taxes that on average are pretty good but in specific cases are wrong. Alternatively, it might be because the tax is simply badly designed or badly conceived. Regardless of the reason the tax is imperfect, the story will be more complex and, as above, it can vary based on expectations and beliefs about the tax in the absence of information. But the overall direction will be that information about a bad tax is more likely to be undesirable than is information about a good tax.

Start with an example similar to the sales tax on food and automobiles considered above but now suppose that it is optimal that they be taxed the same instead of at the actual rates of 2% and 10%, respectively. In this case, the individual without knowledge of the tax and who has the unbiased assumption that they are taxed at a uniform 6% rate is actually assuming the optimal tax and behaving accordingly. Moreover, because expectations are unbiased, the individual's budget balances. Knowledge of the error is unambiguously bad in this case.

The conclusion, however, is not robust: under other assumptions about beliefs it may change even if those beliefs are unbiased (i.e., right on average). For example, suppose the individual believed that the tax rates were flipped – automobiles were taxed at 2% and food at 10%, when the ideal is that they are taxed the same and the actual is that they are taxed at 10% and 2%. We cannot say a priori whether the tax is less efficient at the assumed rates than at the actual rates, even though we can say that both are less efficient than the ideal. Therefore, we cannot say whether knowledge of the tax law increases or reduces the efficiency of the tax.

Overall, with an imperfect linear tax and unbiased expectations, in some cases, knowledge of the tax law is unambiguously bad while in other cases, it is more difficult to know without more structure on the problem. This contrasts sharply with the optimal linear tax and unbiased expectations, where knowledge of the law was desirable.

Example 7 *imperfect labor income tax, unbiased expectations*

As a final example, consider a case taken from Louis Kaplow's paper on accuracy in an income tax. He considers the case where two individuals should be taxed the same but, for unspecified reasons, are taxed differently. For example, if the tax rate on labor income should be, say, 30%, one individual is taxed instead at 25% and the other at 35%. Kaplow assumes that the error is unbiased so that the average is correct but for any particular

individual, it is likely to be wrong. We want to know whether the tax is more efficient with knowledge about the tax rates or without.

The key difference from prior examples is now there are two individuals. Kaplow takes unbiased expectations to mean that people know overall average rates but are wrong about their particular rate. Nevertheless, when it comes time to pay the tax, people will find out they were wrong. Unbiased expectations in this sense, therefore, are what I was calling biased expectations. Any particular individual will be assuming his actual tax payments will be higher or lower than they will really be.

This means that the example is similar to the biased expectation case discussed above in that there are substitution and income effects. If the individuals without information about their actual tax rates assume that they are taxed at the average – 30% – they will initially make choices as if they were taxed optimally. The efficiency gains from the individual taxed at the lower rate are outweighed by the efficiency losses from the individual taxed at the higher rate. The tax is more efficient without specific knowledge of the rates because of this effect.

On the other hand, the individuals will end up paying different amounts than they expect to pay, so they may not have optimized. That is, particular individuals will be wrong about their tax bill even if people are right on average. People will have less money or more money than they expected and if the goods that they would buy change with income, they will not have chosen well. Knowledge of the rates reduces this inefficiency.

The net effect is difficult to predict without more information. It depends on whether the efficiency gains from proper budget balancing (when the individual has knowledge of his particular tax rate) outweigh the efficiency losses from individuals making choices of how much to work based on badly-set tax rates.

A final effect, as discussed above in the case of biased expectation, is that the individuals will update their views on the applicable tax rates, so we would have to adjust our conclusions because of the benefit of this updating. The benefit or cost of lack of knowledge of the law may be limited as people learn their true rates.

Summary

What do we get out of all this? The primary conclusion is that it is very hard to know if and when knowledge of the tax law is desirable. The

three factors isolated here – the type of tax, expectations about the tax, and the quality of the tax – interact in complex ways. Moreover, even if we were comfortable in a given case with how these factors interact in theory, they are hard to observe. For example, it is not easy to determine expectations about a tax because expectations are not directly observable. It is also not easy to determine the quality of a tax as people disagree about which taxes are good ones and which are not. Moreover, subtle differences in hard to observe empirical parameters can change conclusions. And even if we know all of this, we would still be uncertain about how to evaluate the distributive effects that occur within income classes.

Section 3 will consider specific applications of the forgoing discussion in more detail. One overall possible conclusion is that we should limit policies based on whether knowledge of the tax law is desirable to circumstances where we are comfortable with how the factors interact and with empirical estimates of the relevant parameters. An example is the Chetty and Saez experiment discussed above regarding teaching low-income individuals about the decreasing rate structure of the earned income credit.²⁹ Chetty and Saez argue that absent the instruction, individuals believe that their marginal tax rate is simply their average tax rate. Better knowledge of marginal rates in this case is likely to be desirable because marginal rates are below average rates in the relevant range (subject to a caveat that we do not know whether the earned income credit is optimal). Therefore, providing knowledge of the tax law in these circumstances may be welfare improving.

Another example of manipulating access to knowledge of the tax law might be the regulation of tax advisors. Under recent changes to the rules governing tax advice, advice has been made more expensive when the advice is about aggressive tax positions.³⁰ Aggressive transactions take advantage of imperfections in the tax law to produce under-taxation. Knowledge of aggressive transactions is likely to be undesirable. A possible interpretation of these rules, therefore, is that they attempt to limit knowledge of the tax law where it is least valuable.

In other areas, were we are not sure how the factors interact or what the relevant empirics are, we should be caution about designing taxes or making tax policy based on whether knowledge of the tax law is desirable. It is, one

²⁹Chetty and Saez, note __.

³⁰See Brian Gale and David Weisbach, *The Regulation of Tax Advisors*, forthcoming, Tax Notes, March 2011.

might argue, simply too hard to know when knowledge is desirable and when it is not to be confident about using it for policy decisions. For example, we should be cautious about considering salience in tax design. Attempts to set salience optimally are likely no better than random.

Similarly, if a given policy choice has other benefits, we should not let the effects on knowledge of the tax law affect the policy choice. For example, I argued previously that anti-abuse standards such as the economic substance doctrine allow the tax law to be simpler.³¹ A concern with standards, however, is that they make it more difficult to know the precise content of the law; with a rule, you can simply look up the answer while with a standard, you have to guess. If we cannot readily know if and when knowledge of the law is desirable, perhaps this concern should be given little weight.³²

These policy conclusions are speculative. As noted, particular conclusions will depend on details of the circumstances at issue. What is clear, however, is that whether knowledge of the tax law is desirable will depend on a variety of factors and empirical judgments that will be difficult to make.

3 Details

3.1 Tax policy

There are four pieces of tax policy that underlie the discussion above. The key concepts are (1) the measure of efficiency losses from taxation; (2) the reasons for tax law complexity (and hence, the reasons why knowledge of the tax law may be costly); (3) the different types of taxes; and (4) the reasons for focusing on efficiency rather than distributive issues. I discuss each of these, in turn. Our aim is to merely remind readers of the core concepts; more detail can be found in numerous sources.³³

³¹David Weisbach, *Formalism in the Tax Law*, 66 Chicago L. Rev. 860 (1999).

³²This can be taken a step further to argue that they most limit knowledge of the tax law in circumstances when it is least desirable – aggressive transactions. Therefore, the uncertainty that they create is actually desirable.

³³References include Louis Kaplow, *The Theory of Taxation and Public Economics* (2008); Bernard Salanie, *The Economics of Taxation* (2003); *The Handbook of Public Economics* (Alan J. Auerbach and Martin Feldstein, eds.) (Volumes 1 & 2, 1985), (Volumes 3 and 4, 2002), Anthony Atkinson and Joseph Stiglitz, *Lectures in Public Economics* (1980).

3.1.1 The measure of efficiency losses from taxation

If the goal of taxation were just to raise the necessary revenue to run the government, we could simply divide the revenue needs by the number of people in the society and ask each person to pay their pro rata share. Because the tax would not depend on behavior, it would be efficient in that it would not distort decisions. Note that it would reduce people's income (although they would receive public goods and other government services), so people might choose different things than they would without taxation. They might work a different amount, consume different commodities, save a different amount, and so on and so forth. Because most models take the size of government as fixed, these income effects are inevitable and, therefore, are not viewed as inefficient distortions.³⁴

Tax models usually assume that head taxes of this sort are not possible or not desirable. Notwithstanding their simplicity, they may not be desirable because they do not differentiate among people and, therefore, do not redistribute. If we cannot use a head tax, we have to base the tax on observable attributes of the individual, such as how much he consumes, earns, or saves. Any choice of attributes will leave some things untaxed. For example, if we tax consumption, then non-consumption activities such as pure leisure are not taxed. More precisely, if we tax market purchases of consumption, as in a retail sales tax or a VAT, non-market activities – just hanging out – are untaxed. Similarly, if we tax income, leisure remains untaxed. Any tax based on observable attributes other than a head tax will leave some things untaxed.

Suppose one good is not taxed but others are. The untaxed good is usually assumed to be leisure – we don't tax people on the value they get from just hanging out but if they work to earn money to buy consumption goods, we can and do tax their earnings or their purchases. Taxation then introduces distortions as people might change their consumption patterns, or choose to work less (and hang out, untaxed, more), because taxation has changed the relative prices. The distortions come from the change in relative prices – the substitution effect.

To understand the size of the distortion, we can start from a world without taxes and consider what happens when we introduce a tax on a good. The

³⁴An alternative view is that individuals receive tax revenue back in the form of government-provided goods and services so that there is no income effect. Both approaches lead to the same analysis.

price of the taxed good (relative to the untaxed good or goods) will go up, so individuals will reduce their demand for the taxed good.³⁵ As a result, there will be fewer purchases of the good to tax, so tax revenue will decline. People have altered their consumption choices, which hurts them, but there is no offsetting benefit to the government in the form of tax revenue. This difference, the difference between the loss in utility to individual and the government revenue from taxes, is known as deadweight loss. It is the loss from individuals making inefficient choices to avoid taxation.

Deadweight goes up with the square of the tax rate. To understand why, consider the change in deadweight loss from increasing a tax by an increment.³⁶ If the tax were initially zero, the deadweight loss caused by the tax would be very small. To see why, suppose that someone consumes food until he no longer wants to spend any more on food. Say that amount spent is \$1,000. For the very last dollar spent, the benefit of the additional food is just worth the last dollar. The individual decides not to spend an additional dollar more because the benefit of the food is not worth the cost, and decides not to spend an additional dollar less because at that amount, the value of the food is just a tiny bit more than an additional dollar. At \$1,000, the two balance. If taxation causes that choice to be changed to only \$999 spent on food, the utility loss from not having the food but having the dollar is negligible because, at that last dollar, the benefit from spending more on food was negligible. Now if the tax were already positive and we increased it by a further increment, the utility loss would be much greater. If the tax has already distorted food purchases so that only \$900 is spent, going to \$899 imposes real costs. The deadweight loss from additional increments of tax are greater the larger the tax. Correspondingly, if deadweight losses from additional increments to the tax rate get bigger, the deadweight loss itself goes up with the square of the tax rate.

It will be helpful to the discussion below to have the formula for deadweight loss. For a small increase in a tax on good i (starting from zero), we can write:

³⁵Because we are not considering income effects, this will happen for sure.

³⁶This discussion is paraphrased from Jonathan Gruber, *Public Finance and Public Policy* (2005), p. 553. It depends on an assumption that the compensated demand curve is linear. The basic principles do not change with more complex compensated demand curves

$$DWL_{t_i} \approx -\frac{1}{2}t_i^2 \frac{\partial x_i^c}{\partial p_i} \quad (1)$$

where the last term is the change in (compensated) demand for a change in the price of the good.³⁷ This formula is relatively general. It applies to taxes on goods, such as the food example just considered, and to taxes on labor income.

3.1.2 Tax law complexity

If the tax law were a simple head tax or even a linear labor income tax (or equivalently, a uniform commodity tax), it would be very easy for taxpayers to know the content of the law. The tax law must have a sufficient level of complexity that it is plausible that taxpayers do not know the content of the law. The question then arises why it should be complex (retaining the assumption that the law is set optimally). Different reasons for differentiation may lead to different results.

We can divide the tax literature into two central strands: the Ramsey tax literature and the optimal income tax literature. They differ in the assumptions that they make and, therefore, end up imposing somewhat different levels and types of complexity.

The first and older is the Ramsey-tax tradition. It assumes that everyone is the same so it considers a representative individual but rules out head taxes by assuming that at least one consumption good remains untaxed. As noted above, the untaxed good is usually assumed to be leisure. Taxation then introduces distortions as people change their consumption patterns or choose to work less (and hang out, untaxed, more), because taxation has changed the relative prices.

The goal is to choose set of taxes, t_i , to minimize total deadweight loss (the sum of deadweight losses across all commodities) subject to a constraint on the tax revenue the system raises. This maximization problem generates

³⁷The expression in the text assumes that there are no cross-effects – changes in the demand for one good because of a tax on another. In the more general context where there are cross-effects, the expression is similar but includes additional terms to account for the cross-effects.

The expression also assumes that the change in tax rates is from zero. If there is an existing tax and the rate is changed incrementally, the change in deadweight loss will in general be greater than that amount.

a relatively simple formula in which taxes reduce the (compensated) demand for each commodity equally. Under one common formulation of this result, if leisure is not taxed, we want to tax more highly goods that are complements for leisure, in effect trying to indirectly tax leisure.

If fully implemented, a Ramsey tax would be of breathtaking complexity. Commodities under a Ramsey tax are not defined – they are any consumption good. If goods differ in any way, they are different commodities and would get a different tax rate. There would be thousands or hundreds of thousands of different rates. Even the same good at different times of the day might get different rates. Because the models do not have administrative costs, they describe tax systems of essentially unlimited complexity.

One can extend the Ramsey model to include administrative costs.³⁸ For example, we can assume that each additional tax rate imposes additional administrative or compliance costs and then calculate the optimal rate structure. Unless administrative costs from additional rates were very high, however, the tax might still be complex. Note that an optimal Ramsey tax with administrative costs would, in the language used above, be an imperfect tax, such as that considered in Example 5.

The second tradition, originating with James Mirrlees, focuses on the distributive basis for taxation.³⁹ In Mirrlees’s model, people vary by their ability to earn wage income but the government cannot directly observe ability; it can only observe total wage income. The goal is to maximize a social welfare function. Because marginal utility declines with income, social welfare can be improved through redistribution, so we do not want a uniform head tax.⁴⁰ We cannot directly observe ability, however, so we also cannot have a differentiated head tax. Therefore, we tax wages as an indirect method of taxing ability. That is, unlike the Ramsey model which arbitrarily imposes a restriction in available taxes, the Mirrlees model imposes restrictions based on the information available to the government and an explicitly stated social goal.

The problem with taxing wages rather than innate earning ability is that if the government taxes high earners at a high rate, they can reduce their

³⁸See Yitzhaki, AER, and Wilson, ___.

³⁹Mirrlees (1971).

⁴⁰Mirrlees assumes that marginal utility declines with income. If we have an individualistic social welfare function, redistribution can increase social welfare because a unit of consumption to a less-well off person will generate more utility than the same unit to a well-off person.

effort and earn less, in effect pretending to be low-ability individuals. The optimal solution, which is complex, trades off the benefit of redistribution to low-ability individuals and the costs of higher-ability individuals reducing their labor effort.⁴¹

The core of the Mirrlees model considers only labor income taxes. The rate structure may be complex and individuals may not understand the structure. Moreover, extended versions of the model add additional complexity.⁴² One reason is that it may be desirable to tax more highly goods that are complementary to leisure to make reducing labor effort less attractive. The result may be a tax which differentiates among goods in a manner not too different from a Ramsey tax. Another reason is that taxpayers may have preferences related to their underlying ability. The tax system might be able to take advantage of these preferences to better tax ability, such as, for example, by taxing goods desired by those with high abilities (regardless of their incomes). And taxpayer heterogeneity (holding income fixed) might optimally be taken into account. Thus, if some get less utility from a given income than others (say some people are disabled and need additional income to attain a given level of utility), the tax system might optimally differentiate among taxpayers with different characteristics.⁴³

Regardless of which approach is taken, the Ramsey approach or the Mirrlees approach, the optimal tax system is likely complex. Knowledge of the tax law, even when the system is set optimally, is an issue. Moreover, suboptimal systems, such as current law, are likely to be complex for a wide variety of reasons. For example, if the tax system attempts to target higher or lower taxes on a particular group, specifying the group, particularly when people can change their behavior to be in or out of the group, may be difficult and require complex rules.

⁴¹To get an intuition for the result, imagine that the tax rates are set optimally and we imagine tweaking the rate at some income level up by an increment. There are two effects. First, individuals at that income level will face a higher marginal tax rate. This reduces their labor effort, creating a social cost. Second, individuals at a higher income level will pay more tax, but their marginal rates will stay the same. This creates a distributive gain (people above the income level pay the additional tax; people below do not). These two effects will offset at the optimum.

⁴²See Louis Kaplow, *The Theory of Taxation and Public Economics* (2008) p. 135.

⁴³Cite disability paper

3.1.3 Types of taxes

A central distinction in the examples discussed above is between Pigouvian taxes and other taxes. A Pigouvian tax or subsidy is designed to cause people to internalize externalities. It is optimally set equal to the marginal harm or benefit from an activity. For example, a carbon tax would be set equal to the additional harm from climate change due to another unit of carbon in the atmosphere. We can think of Pigouvian taxes as substitutes for tort liability. Both are systems for causing individuals to consider external harms when they take actions.

Other taxes are designed either to simply raise revenue for the purchase of public goods or to redistribute. In the Ramsey tax tradition, there is a fixed government budget that must be met. In the Mirrlees tradition, taxes also redistribute (and depending on the model, some portion may be spent on public goods).

As noted, the effect of knowledge of the tax law may vary depending on the type of tax. Considerations about knowledge of tort laws likely apply to Pigouvian taxes but not to other taxes. Therefore, understanding the type of tax at issue will be important to understanding the effect of knowledge of the tax law.

Unfortunately, it is not always easy to determine whether a particular tax is a Pigouvian tax. Consider, for example, a tax on complements to leisure (and subsidy for substitutes) imposed with the goal of reducing the labor/leisure distortion created by an income tax. The tax on complements to leisure can be seen as a Pigouvian tax on the fiscal externality caused by individuals working less. We would want people to know about this tax because we want it to alter behavior – to cause people to work more.

3.1.4 Efficiency v. distribution

A final aspect of the examples above is that they only considered the efficiency aspects of knowledge of the tax law, not the distributive effects. Policies that affect knowledge of the law might have distributive effects, so perhaps these should also be taken into account.

The reason I did not is that we can always adjust tax rates to offset the distributive effects.⁴⁴ For example, suppose that knowledge of the tax

⁴⁴For a general explanation of this approach, see Louis Kaplow, *The Theory of Taxation and Public Economics* (2008).

law allows high income taxpayers to reduce their marginal rate from 35% to 30% and we are considering a policy that would reduce knowledge of the tax law and thereby raise their rate to 35%. To determine whether the policy is desirable, we want to compare it to alternative policies. For example, instead of adopting the policy concerning knowledge of the law, we might simply increase marginal rates so that the effective rate goes up to 35%. The two policies – the knowledge policy and the increase in rates – have the same distributive effect but likely have different efficiency effects. We can compare them simply by examining their efficiency effects. Alternatively, we could compare not adopting the knowledge policy and adopting it plus lowering rates to 30%. Again, the distributive effect would be the same and we need only examine the efficiency effects.

For this reason, the focus on the discussion was on the efficiency effects of knowledge of the tax law. With a sufficiently flexible rate structure, this is all we need to worry about.

3.2 Deadweight loss with imperfect optimization

The examples in Part 1 examined what happens if taxpayers do not know the full content of the tax law. The deadweight loss measures discussed immediately above, however, are constructed assuming full optimization. We must confirm that the intuitions about deadweight loss carryover to the case where taxpayers fail to optimize fully because of lack of knowledge.

Fortunately, a recent paper by Raj Chetty, Adam Looney, and Kory Kroft confirms this.⁴⁵ The key idea in their argument is that for a given amount of knowledge of the tax law (salience in their case), we can observe actual behavior, even if this is not fully optimizing behavior. For example, if the tax law is not salient, people may not adjust their purchases as much as they otherwise would. Given that effect on behavior, we can measure the utility losses using the standard measures.⁴⁶ When this is modeled, the result is

⁴⁵Raj Chetty, Adam Looney, and Kory Kroft, *Salience and Taxation: Theory and Evidence*, 99 *American Economic Review* 1145 (2009)

⁴⁶A key difference that makes the salience case more difficult than the lack of knowledge case, is that Chetty, Looney and Kroft are assuming that lack of salience results from a failure to fully optimize. It is a behavioral problem. They then must recover an underlying rational utility function that they can use to measure utility losses. They do this by making the assumption that adjustments to fully salient price changes are optimal. (See the Appendix for a discussion.) In the knowledge case, knowledge of the law is likely to be

deadweight loss is related to perceived taxes. (A more complete discussion of their argument is contained in the Appendix.)

To illustrate, suppose that the actual tax on a particular good i is t_i but individuals misperceive this and instead think that the rate is $\theta_i t_i$ where θ_i is a fraction. The θ_i can be higher or lower than 1, so that people think some taxes are too high and some are too low. Chetty, Looney and Kroft show that we can write the formula for deadweight loss as:

$$DWL_{t_i} \approx -\frac{1}{2} (t_i \theta_i^c)^2 \frac{\partial x_i^c}{\partial p_i} \quad (2)$$

As can be seen, the formula is same as the previous formula for deadweight loss (Equation 1) except that the tax rate t_i is replaced with $\theta_i t_i$.⁴⁷

When there are no income effects, deadweight loss is a function of the perceived tax rate. The basic idea is intuitive: people adjust their behavior based on their perceptions even if those perceptions are false. If they perceive the tax rate on an item is higher (or lower) than it actually is, they will react accordingly.

Equation (2) allows for income effects, which can make the analysis more complicated. To illustrate, suppose that a taxpayer (incorrectly) thought that for a particular commodity, the tax rate was very low. He would not adjust his purchases of the good very much because of this view. The substitution effect would be correspondingly smaller and the deadweight loss less. The individual, however, has to pay tax at the correct rate, which means that he may not have optimized. For example, suppose that the tax is large, such as a substantial sales tax on a big purchase such as an automobile. If the individual underestimates the tax and therefore does not properly adjust his automobile purchases, he may not have enough money for other things he wants. Had he correctly estimated the tax, he might have adjusted his purchases. Thus, if there are income effects, lower θ_i 's may or may not lower deadweight loss.

Note, however, in the case of unbiased expectations, the θ_i 's average to 1, so there are no income effects. People correctly estimate their average tax rates so they correctly estimate their budgets. There should be no income effects, and deadweight loss move with the variance of the θ_i . This is why,

expensive, so failure to obtain knowledge can be fully rational. We need not worry about how to construct a welfare measure when people are not optimizing.

⁴⁷Mention compensation.

in the case of optimal linear taxes, knowledge of the tax law is desirable – it reduces the variance of the θ_i .

4 Applications and Extensions

4.1 The incentive to seek knowledge of the tax law

An alternative approach to understanding knowledge of the tax law is to ask whether the right amount will be supplied by the market. That is, do people have a socially optimal incentive to seek tax advice? I start with the case where the tax is linear and is set optimally, and then consider other cases.

Suppose that the tax is set optimally and consider an individual seeking advice on the tax law. Does the individual's private value of the advice equal the social value of the advice?

An initial intuition might be that taxpayers have too low an incentive to seek advice. If tax advice reduces deadweight loss, taxpayers may have too low an incentive to seek advice because they only consider the effect of advice on their own utility. Reducing deadweight loss is, in a sense, might be like a positive externality from advice.

The initial intuition may not be correct, however. Deadweight loss is the difference between the individual utility losses from taxation and the revenue raised. Individuals deciding to seek tax advice will fully consider the effects on their utility. The only possible externality from seeking advice, therefore, is the change in revenue. In our setting, where taxpayer views about the content of the law are unbiased, taxpayers may equally end up paying more or paying fewer taxes once they learn the actual content of the law.

To illustrate, suppose that there are only two goods, food and automobiles and individuals purchase equal amounts of each. The tax rate on food is 2% and the tax on automobiles is 12%, for an average of 6%. If taxpayers do not know the actual tax rates but have unbiased beliefs, they can believe the actual rates are any pair of numbers that average to 6%. They may believe that the tax on food is higher than it is and the tax on cars lower, or vice versa.

It is not clear if they learn the true rates whether they would pay less or more tax. If the perceived tax on food is higher than the true tax and the tax on automobiles lower, when individuals find out the true rate they will increase their food purchases (the low-taxed item) and reduce their automobile

purchases (the high taxed item). Whether this reduces revenue depends on the demand elasticities for each of the items. If their perceptions are reversed (they believe tax on food is lower than it really is and tax on automobiles is higher than it really is), the results may be the reverse. Therefore, incentives to seek advice may be too low or too high.

This conclusion contrasts with the Louis Kaplow's argument that when a tax is inaccurate, there tends to be a socially excessive incentive to seek advice. Consider a tax that should treat two individual at the same rate but for some reason does not. Assume taxpayers are risk neutral, do not know the exact tax rate, and have unbiased expectations about the rate. Before advice, they would assume that they are taxed equally. After receiving advice, some individuals will know that they are over-taxed and some under-taxed. Because deadweight loss goes up with square of tax rate, there is greater deadweight loss when taxpayers have correct information about an inaccurate tax. Therefore, taxpayers have too high an incentive to seek advice.

The two results – that in the current setting, taxpayers may either have too high or too low an incentive to seek advice and in Kaplow's setting, they tend to have too high an incentive – can be reconciled by examining the assumptions each analysis makes. In our case, we are assuming that the tax law is set optimally but taxpayers are misinformed about its contents. Advice in this setting is valuable and the incentive to seek advice may be correct or even too low. In Kaplow's case, the law is set in error, but is unbiased. Without advice, taxpayers use the unbiased expectation, which is better than if, with advice, they use the incorrectly set rates. One way to summarize the difference is that learning about an accurate tax is may be good thing, but learning about an inaccurate tax is often a bad thing.

4.2 The regulation of advisors

One of the most significant areas of tax regulation in the last decade has been the regulation of tax advisors. Gale and Weisbach (draft) provide a detailed description of these changes. The one sentence description is that tax advice is more expensive to give and its value lower. In a sense, the government has raised the cost of obtaining knowledge of the tax law. The question is whether this set of regulations is desirable.

Evaluating these regulations is complex and raises issues beyond merely whether knowledge of the tax law is desirable. Tax advice takes place in a

complex institutional setting that involves law firms, clients (which are often large entities), and the government. Principal/agent problems abound in each of these settings and interact when tax advice is sought and given. Tax laws are imperfect (and sometimes deeply flawed) and constantly changing. Tax rules interact with other regulatory regimes, such as the accounting rules or securities laws. There are multiple purposes behind tax laws, and tax laws may often substitute for simple direct spending. Government auditing of tax returns is imperfect with penalties too low to make up for this, creating incentives to play the audit lottery, which in turn affects the advice taxpayers seek. There are multiple layers of regulation of legal advice. And all of this takes place within the context of democratic law making with the attendant free speech concerns.

Given these complexities, I cannot do more than suggest some directions for investigation here. In particular, we might expect, as a first cut, that the regulation of tax advice should increase its cost in cases where knowledge of the tax law is undesirable and reduce its cost where knowledge of the tax is desirable. Given the subtleties of the relevant distinctions – relating to expectations, quality of the tax, and the kind of tax – this will not be easy.

Nevertheless, one way to view the regulation of advisors is exactly this – they try to limit access to information about the tax law when it is undesirable.⁴⁸ Consider the following new rules.

- Circular 230 § 10.35. Circular 230, which governs attorneys and other authorized individuals practicing before the Service, was significantly expanded in recent years and now regulates a class of written tax advice known as “covered opinions.”⁴⁹ Covered opinions are essentially tax opinions relating to tax shelters or transactions with significant tax avoidance potential. Under these new rules, the content and form of advice is heavily regulated. They impose general diligence requirements and scope requirements making access to advice on these transactions more expensive.
- Strengthened Requirements for Taxpayers Seeking Penalty Protection under §§ 6662, 6662A, and 6664. Taxpayers can generally avoid the various understatement penalties by relying on tax advice.⁵⁰ The rules for

⁴⁸Cite to Claire’s paper in Virg. Tax Review on disrupting the market for tax advice.

⁴⁹31 U.S.C. §330 (2006) and 31 CFR § 10.0-93 (2003)

⁵⁰See §§ 6664 and 6664A.

such reliance have undergone significant amendments in recent years. The circumstances under which taxpayers may claim reasonable cause and good faith have been markedly reduced, especially in situations involving so-called “reportable transaction” understatements which are understatements related to aggressive transactions. Moreover, advice from many advisors—including those involved in structuring parts of the transaction—may not be relied upon.

- Expanded List-Keeping and Disclosure Requirements under §§ 6111 and 6112. Under current law, the record-keeping requirements on advisors are much more onerous, and penalties for violations are substantially greater than under prior law. Material advisors now are required to disclose client participation in reportable transactions and keep lists of client involvement. Under prior law, the duties imposed on advisors only pertained to tax shelter transactions, a much more narrowly defined class. Additionally, penalties for noncompliance were increased substantially. In one recent example, KPMG alone was forced to pay \$100 million in civil fines for failure to register the tax shelter transactions with the Service, as part of its deferred prosecution agreement with the government. These changes have further increased the cost of advice to taxpayers, as advisors demand higher compensation for the increased burdens and potential liability they now face.

All of these changes can be seen as attempt to identify situations where knowledge of the tax law is undesirable and to increase the costs of advice in these situations. In particular, if we think of tax shelters as unintended interactions of the tax rules that create gaps, shelters likely fit into the class of cases where knowledge is undesirable.⁵¹ They are cases where the law is likely suboptimal. While other factors could cause knowledge of the tax law in these cases to be desirable, as a first cut we might suspect that it is not. Therefore, one way to see this set of regulations is as an attempt by the government to limit or make more expensive access to knowledge of the tax law where such knowledge is undesirable.

⁵¹Reference definition of shelter.

4.3 Manipulating Salience

Imagine a tax that is completely hidden: taxpayers are given no information about the tax until the last day of the year, at which point it is announced (so that they may easily fill in their tax returns). To prevent the government from acting strategically, we could force the government to write the tax down on the first day of the year but only reveal it at the end of the year.⁵² Would this tax be a good idea or a bad idea?

This somewhat crazy example is simply an extreme version of more modest proposals aimed at manipulating knowledge or salience of the tax law. For example, Raj Chetty and Emmanuel Saez conduct an experiment in which they teach low income taxpayers about the earned income tax credit, manipulating knowledge of the tax law by increasing it.⁵³ They suggest that individuals confuse their average and marginal rates. Marginal rates decline in the relevant range, which means that individuals are assuming that marginal rates are higher than they actually are. Increasing knowledge of the law, they suggest, can reduce deadweight loss.

Similarly, Jeffery Liebman and Richard Zeckhauser suggest that lack of knowledge can be a good thing if, when rates are increasing, people assume that their marginal rate is equal to the average rate. They calculate that deadweight loss from taxation is roughly half because of this lack of knowledge of marginal rates. The implication is that we should not make the marginal rate structure more transparent and might even benefit by making it less visible.

While Chetty, Looney and Kroft are careful not to say this specifically, a casual reading of their paper might lead one to conclude that they advocate using low-salience sales taxes to reduce the deadweight loss of commodity taxation. They argue that people might systematically underestimate sales taxes and in the absence of income effects, this might reduce deadweight loss.

Deborah Schenk suggests using low salience taxes to overcome political breakdowns. In particular, if the legislature cannot keep taxes high enough to support the desired expenditures, she suggests that they might use low-

⁵²As noted in note __, Saul Levmore suggested this example to me many years ago.

⁵³Chetty and Saez, note __, p. 34. (“the results suggest, for instance, that high phase-in rates coupled with low phaseout rates over a long range could maximize the impact of the EITC on work effort. Such a structure would make the benefits of working in the phase-in region most salient, and may permit the work disincentives in the phaseout region to be framed as negligible.”)

salience taxes to be able to raise the necessary revenue. Ed McCaffery demurs, however, arguing that legislatures use low-salience taxes to pursue illegitimate ends.⁵⁴

The suggestion above was that that it is sufficiently difficult to know when knowledge of the tax law is desirable, that we should be cautious about attempting to manipulate salience. Doing so might very well increase rather than reduce the costs of taxation. The question I address here is whether we can be more precise: what are the factors that might help tell us when good idea and when not?

As a primary matter, we should only manipulate salience when we it is likely that it will increase efficiency.⁵⁵ To know when reducing (or increasing) salience increases efficiency we have to determine expectations about the tax in the absence of good knowledge, the quality of the tax (something likely to be disputed in most cases) and the type of tax. Expectations about the tax in the absence of knowledge can be determined through empirical or perhaps experimental studies such as those cited above. The quality of most taxes, however, is likely to be disputed. There is little consensus about which taxes are desirable. Even the type of tax may not be easy to determine. For example, a tax on complements to leisure – suggested by the standard Ramsey formulation of the optimal revenue raising tax – can be seen as Pigouvian because its goal is to reduce the fiscal externalities caused by working less (due to a tax on labor income).

This is not to say that we never know when knowledge of the tax is undesirable. For example, Liebman and Zeckhauser consider graduated labor income taxes and argue that lack of knowledge about marginal rates reduce the deadweight loss. They claim that individuals assume the marginal rate is the average rate, so expectations are unbiased and there are no income effects. If this is true (and marginal rates are increasing), it may not matter that the rate schedule is not likely to be optimal. For any given rate, the labor/leisure distortion is less.

A current and real world application of this logic involves a set of provisions that raise marginal rates within certain (high) income ranges by phasing out exemptions and deductions. These provisions, known as PEP and Pease, are not transparent, and most affected taxpayers likely do not know about

⁵⁴Edward J. McCaffery, *Cognitive Theory and Tax*, 41 UCLA Law Review 1861 (1993).

⁵⁵Even if we were able to determine that reducing knowledge of a given tax increases efficiency, we would still have to believe that the resulting loosening of the political restrictions on taxation are desirable. Compare Schenk, note __ and McCaffery, note __.

them.⁵⁶ They have been criticized as hidden taxes, and commentators have suggested that they should be repealed for this reason.⁵⁷ The question is whether they are more or less desirable because they are hidden.

Because they raise marginal rates within a certain income range, an initial thought is that they resemble the Liebman Zeckhauser case in which people knowing their average rate will underestimate their marginal rate, increasing efficiency. On the other hand, marginal rates decline once the phase-outs are complete, so other people may overestimate marginal rates. The net effect could go either way. They were repealed in 2001 as part of the Bush tax cuts but are scheduled to be reinstated when those cuts expire. It is not clear whether this is desirable.

Finally, consider a tax that (I believe) is suboptimal and likely simply bad: the tax on capital income. Is knowledge about the details of the capital income tax (which consist of very large portions of the Internal Revenue Code) desirable? On the one hand, all else equal, knowledge of a bad tax is less likely to be desirable than knowledge of a good tax. On the other hand, the current tax on capital is highly incomplete and can be avoided. Knowledge of these gaps might reduce the tax which is a good thing. The question is whether the shifts inefficiencies in investment patterns needed to avoid the tax create greater social losses than the tax. We cannot know this without substantially better models of capital income taxation.

4.4 Simultaneously optimizing tax and information

To be supplied.

5 Conclusion

I explored here whether and when knowledge of the tax law is desirable. It turns on the interaction of at least three hard to observe factors: taxpayer expectations, the type of tax, and the quality of the tax. Notwithstanding

⁵⁶PEP stands for Personal Exemption Phase-out and was enacted in 1990 but was based on a related provision that was enacted as part of the 1986 Tax Reform Act. See section 151(d)(3). Pease, which phases out itemized deductions, is named after Congressman Donald Pease, was also enacted in 1990. See section 68.

⁵⁷See Statement of Richard M. Lipton on behalf of the American Bar Association Section of Taxation before the Committee on Finance of the United States Senate, April 26, 2001, found at <http://www.abanet.org/tax/pubpolicy/2001/tes010426simpl/home.html#toc>

the complexity of the inquiry, it will be central to a number of policy questions. I explored two such questions here: whether taxpayers have the correct incentives to seek knowledge of the tax law and how tax advisors should be regulated.

6 Appendix with details on Chetty, Looney and Kroft

They argue that consumers consistently underestimate sales taxes (which are imposed at the cash register and are not stated in the price of the good) and then analyze the welfare effects. Although their model is a behavioral economics model, created to measure welfare when consumers fail to optimize, we will argue that it can be used to generate intuitions where information is costly and consumers properly optimize given imperfect information. We describe the set up and conclusions of their model and then show how it applies in our case.

Set-up:

Two goods: one is taxed, one untaxed. No preexisting taxes.

Utility is additively separable in the goods – no cross-priced elasticity.

Firms are price takers, constant returns to scale so that the pretax price of the taxed good is fixed.

With full optimization (perfectly salient taxes), changing either the price or the tax rate of the good would lead to the same change in behavior. But if taxes are not fully salient, a change in the tax rate will cause a different change in behavior than the change in the price.

Key variable = θ_i = salience of tax i = percent change in demand for a good b/c of change in the tax divided by percent change in demand because of change in price. Similarly define compensated $\theta_i = -$ the ratio of the income-compensated tax effect and the normal compensated price effect.

If $\theta_i = 1$, know the law perfectly. If $\theta_i = 0$, assume tax = 0%. If $\theta_i > 1$, people respond to the tax more than to the price – they overestimate the tax rate.

Chetty assumes that $\theta_i < 1$ because of behavioral effects. We assume that average $\theta_i = 1$ but any given $\theta_i \neq 1$ because we assume no systematic bias. As we will see, this drives the conclusions.

Want to measure losses due to taxation

Defined as how much could have been raised in revenue if we were to use a lump sum tax instead of the distortionary tax, holding utility constant. It is the total value of the purchases that fail to take place because of the tax.

CLK deal with the case where individuals fail to optimize. Have to make some assumptions about behavior and utility to be able to calculate DWL in this case. They make two assumptions – we examine them to make sure that they apply in the case where there is an information cost to determining the content of the law.

A2: when taxes are fully salient, people optimize.

Straightforwardly applies to our case.

A1: taxes affect utility only through their effect on consumption.

Trickier. Violated if cognitive cost of taxation because then they affect utility by more than their effect on consumption. The cognitive costs reduce utility. Here, we are assuming information about the content of the law is costly. Therefore, there may be information costs to determining the content of the tax law that would not arise for an equivalent price increase. Therefore, A1 arguably does not apply.

In section 3, however we are just trying to determine whether increasing information about the content of the law increases or decreases DWL. Not yet considering the decision to acquire information. Will come back to that and revisit assumption A1. It will turn out that the DWL formula discussed here still works once we add information costs – it turns out to be DWL net of any information costs. Therefore, A1 is sufficient for purposes of the present discussion.

CLK Results:

Normal – DWL goes up with the square of the tax rate and depends on the compensated elasticity of demand. (Note that this ignores cross-effects, which are assumed away because of the choice of utility function.)

With salience, get the same formula except replace t with $-$ the salience adjusted tax rate. We get

If no income effect, lower salience is better.

All there is, is substitution effect and lower salience lowers this. As if the tax rate were lower. You pay the tax but it has no behavioral effect. With no income effect, having less income because you pay the tax doesn't affect choices.

If there are income effects, however, lowering salience may not lower DWL. The reason is that if people do not pay attention to taxes, they may not balance their budget correctly.

They still have to pay the tax, after all. If they ignore it when making (some) budgeting decisions, may end up with suboptimal bundle of goods. So even if salience is zero – people initially act as if taxes were zero, might have high DWL.