Measuring Tax Complexity

Dominic Kassirra, Ryan Oprea, Alex Rees-Jones, Zamir Ticknor

Income taxes are often lamented for their complexity. This tax complexity can burden taxpayers. For example, more complex taxes may lead filers to spend more time and effort calculating their annual tax liability or complying with tax reporting requirements. More complex taxes may also lead citizens to be more confused about tax incentives, causing them to fail to adapt their behavior to respond to financial incentives.

Despite the importance of tax complexity, few formal and systematic measures of tax complexity are currently available. This absence is not due to a lack of interest, but instead largely due to the difficulty in formally defining "complexity" in a manner that is amenable to measurement. Existing methods include counting the number of words in a tax provision, readability scores of tax provisions, or analyzing user questions in online forums. These methods provide valuable signals of complexity, but are ill-suited for the type of precise, individualized measurements that would be most useful for analysis.

We propose an approach to measure tax complexity that builds on notions of complexity in the computer science literature. One means of assessing the complexity of a given operation is assessing the number of discrete cells of information that must be accessed. Applying that notion to the tax context, we assess the number of cells of the tax form that tax filers must attend to in order to accurately forecast their tax level or tax rate. When taxpayers can correctly determine these key tax variables while attending to only a few cells of information—say, by attending to only their wage income—we say that taxes are comparatively simple. If the taxpayer must attend to many individual items for an accurate forecast—say, wage income, capital gains, business income, and itemized deductions—we say that taxes are comparatively complex. This notion allows us to capture changes in average complexity of the tax code over time, but also to capture differences across people in the complexity inherent in their tax situation at the same moment under the same tax code. This conceptual approach can be formally pursued by applying a variant of the "sparsemax" approach of Gabaix (2014).

Our approach can generally be applied to specify the complexity of administrative forms with repeated cross-sectional data. As a concrete application, we apply this approach to readily accessible IRS Tax Model Files of the years 1966–2014 using the NBER taxpuf program, a more elaborate version of NBER TaxSim.

With our approach, we can answer the following important questions about tax complexity in the United States:

1. For the average tax filer, how complex is the current U.S. tax system?
2. How is complexity different across different filers? Does the burden of tax complexity fall unequally, and if so, who bears the burden?
3. How has complexity evolved across time and across groups? How have the tax reforms of the past 50 years changed the distribution of tax complexity's burden?

We document that tax complexity increased on average over time, in line with anecdotal evidence in the literature. We observe a notable jump in tax complexity with the implementation of the Tax

Reform Act of 1986, and document other jumps in tax complexity around major U.S. tax reforms. Notably, high earners saw a notable decrease in tax complexity over time, while the taxes of low-income groups became increasingly complex. A key driver is the increase of refundable credits in the U.S. tax system over time. One such example is the EITC, which expanded in magnitude and eligibility. We document that for the vast majority of low-income earners, wages and salaries were sufficient to get an accurate estimate of their MTR in the early sample period. But, after the introduction of the EITC in 1975, many low-income earners now had to attend to at least two items, wages/salaries and the EITC, which substantially increased complexity.

Another important application of our measure is to estimate or control for the effect of tax complexity on economic behavior. Public finance researchers are aware that changes in tax codes often result in changes to both incentives and complexity. If researchers use historical changes in incentives to study behavior, such as labor-supply elasticities, and complexity also affects that behavior, omitting complexity induces bias. Similarly, if researchers rely on cross-sectional variation in tax incentives, cross-sectional differences in tax complexity can bias results. Utilizing our approach or the results we provide, researchers can control for changes and differences in tax complexity, removing the bias caused by tax complexity differences.

Acknowledgment: This research was supported by the National Science Foundation under Grant SES-1949366.