The Role of General Equilibrium Models in Tax Policy Analysis

Florenz Plassmann

THERE IS a long tradition of economic analysis to guide tax policy. In 1756, François Quesnay suggested that taxes on labour or on output reduce a nation's wealth while a tax on land has no such effect. He therefore recommended that governments should raise their revenues from land, so as to reduce the nation's wealth as little as possible.[1] Twenty years later, Adam Smith reached a similar conclusion, saying that the rent of land was the source of revenue that can best bear a peculiar tax, because "[t]he annual produce of the land and labour of the society, the real wealth and revenue of the great body of the people, might be the same after such a tax as before."[2] Until the late 19th century, most economists agreed that taxes on land have a different effect than taxes on labour and capital. Neoclassical economists, however, have tended to treat land as a form of capital and differentiate only between labour and capital. As a result of aggregating factors with different properties into a single factor "capital", most existing neoclassical models of taxation are unlikely to correctly determine the economic distortions that are caused by taxes on labour and capital, and those models are also incapable of estimating the potential benefits of shifting taxes from labour and capital to land. This report reviews the differences between various kinds of taxes and to summarise the most important properties that an economic model for tax analysis should possess.

Dr **FLORENZ PLASSMANN** *is Assistant Professor of Economics at the State University of New York (Binghamton)*

The excess burden of taxation

ECONOMISTS HAVE collected convincing evidence that people respond to changes in relative prices, and that, if taxation leads to changes in relative prices, people will respond by shifting their activities to sectors that are not taxed.

Economists recognise that this effect of taxation can be desirable if an activity generates a negative externality.[3] The price of gasoline, for example, reflects only the cost of producing and selling gasoline, but it does not reflect the social cost of the pollution that is caused by cars. Imposing a tax on gasoline can raise the price of gasoline to include the cost of pollution and thereby force consumers to take this social cost into account. Consumers will respond to the price increase by driving less. If, however, an alternative fuel that does not generate any pollution is discovered, the price of fuel without any tax will lead to the socially optimal consumption of fuel. No tax should be imposed on that fuel if the sole purpose of a tax on gasoline is to reduce inefficiently high gasoline consumption to the socially efficient amount, rather than to generate government revenue.

If an activity is not socially harmful, then any tax that raises its market price will

reduce the activity to a socially inefficient level. Economists call the social cost of this reduction in taxed activities "excess burden." Excess burden is measured as the lost benefit that would have accrued to people from the activities that they would have undertaken in the absence of the tax. It is straightforward to show that the excess burden is roughly proportional to the square of the tax rate, which implies that doubling the tax rate will quadruple the excess burden.[4] The exact size of the excess burden depends on the elasticities of the demand for and the supply of the activity, the tax rate, and the total size of the taxed sector. The less elastic the demand and the supply schedules, the smaller is the excess burden, and vice versa. If either the demand or the supply schedule is perfectly inelastic, then the excess burden is zero.

All taxes that are imposed on activities with elastic demand and supply schedules have excess burdens. A tax on labour income has an excess burden, because such a tax raises the cost of labour relative to the cost of leisure. To the extent that a tax on *labour* income increases the cost of employment, it provides an incentive to firms to hire fewer workers. To the extent that a tax on labour income lowers the net wage that workers receive, it provides an incentive to work less and to consume more leisure. To the extent that workers are mobile, a tax on labour income provides an incentive to migrate to areas without or with lower taxes on labour income. Because work is not a socially harmful activity, its reduction constitutes an excess burden.

A tax on capital has an excess burden, because such a tax lowers the return that owners receive from capital. Because new capital is financed through savings, a tax on capital lowers the reward from saving relative to the reward that a person receives from consumption, and provides an incentive to consume a larger share of one's income. Because saving is not a socially harmful activity, its reduction constitutes an excess burden. To the extent that some capital is mobile (for example, machinery), a tax on capital provides an incentive to move mobile capital to areas that do not impose taxes on capital. Whether or not capital is mobile, a tax on capital reduces the incentive to replace capital as it depreciates. Because the presence and use of capital are generally not socially harmful, the reduction of the available capital constitutes an excess burden.

Indirect taxes also have excess burdens. Excise taxes (taxes on goods like telephone services, air travel, and luxuries) raise the costs of these goods relative to others and induce consumers to consume inefficiently small quantities of these goods. Sales taxes raise the price of consumption relative to the price of saving, and lead to inefficiently low levels of consumption. Taxes on foreign imports raise the price of foreign goods relative to the price of domestic goods, and lead to an inefficiently low level of consumption of foreign goods. To the extent that the consumption of domestic and foreign goods is not socially harmful, their reductions constitute an excess burden.

Taxes that are imposed on goods with perfectly inelastic demand or supply schedules do not have excess burdens. If a good's demand schedule is perfectly inelastic, a tax increases the market price of the good by an amount equal to the tax, and the quantity exchanged remains unchanged. However, because all consumers face budget constraints, no good can have a demand schedule that remains perfectly inelastic for all price increases, so even a tax on a good with perfectly inelastic demand (for example, heart-transplants) will have an excess burden if it increases the price to a level at which consumers cannot afford the good anymore.

If a good's supply schedule is inelastic, a tax reduces the income that the good's owner receives, but the tax leaves the market price as well as the quantity supplied unchanged. Only land and non-exhaustible natural resources have completely inelastic supply schedules, because their supplies can neither be increased through production nor decreased through depreciation. Sometimes it is argued that a poll tax does not have an excess burden, but this is correct only if labour is assumed

to be completely immobile, completely insensitive to the tax with respect to decisions about whether to have children and whether to commit suicide, and never so poor as to be unable to afford the tax.[5] Economists generally agree that the only taxes without excess burdens are taxes that compensate for externalities (these actually improve economic efficiency), taxes on land and other inexhaustible natural resources, and taxes on exhaustible natural resources that are levied in such a way that the total value of taxes collected is independent of decisions about the rate of exhaustion. If a government wants to raise revenue with as little harm to the economy as possible, these are the taxes that are attractive.[6]

Applied general equilibrium models for the analysis of the excess burden of taxation

THERE ARE two ways to examine the economic effects of taxation. If the tax has already been introduced, one can undertake an econometric analysis of the situations before and after the tax has been introduced to determine the effects of the tax. This, of course, is not possible if the tax has not been introduced yet, and if one wants to estimate the economic effects of introducing the tax in the future. Because taxes can distort an economy in many different ways, economists have begun to use general equilibrium models to study the economic effects of planned or hypothetical tax programs, now that computing costs have fallen as low as they have. A general equilibrium model describes the circular flow of input factors and output products between consumers and producers, and, as opposed to a partial equilibrium model, emphasises that economic actions may have impacts on relative prices throughout the economy, now and in the future.[7] A solution to a general equilibrium model consists of a set of prices for all goods that ensures that, in all markets, quantity supplied equals quantity demanded. Only under very restrictive model assumptions it is it possible to find this set of prices with analytic methods. Models with numerous sectors, agents, and regions are usually too complex to have analytic solutions, and they need to be solved numerically. Such models are called applied general equilibrium (AGE) models. AGE models were added to the economists' toolkit in the 1970s, and have found widespread use since.[8]

The use of AGE models for policy analysis requires a trade-off between complexity and expected precision. To keep the model manageable, it is necessary to aggregate parts of the economy (for example, one might aggregate all food processing industries into a single "food industry", or all consumers into a single "representative consumer"). The higher the level of aggregation, the easier it is to solve the model, but the less precise are the predictions of the model, because aggregate models describe only the average development of the aggregate units, which can be imprecise if the units that were aggregated are very heterogeneous. The lowest *possible* level of aggregation is determined by the available data that can be used to calibrate the model;[9] if only highly aggregate data are available, it would be meaningless to generate a very disaggregate model, because it would be impossible to calibrate such a model with the desired degree of precision.

The simplest (that is, most aggregate) general equilibrium model covers a single time period, and consists of a single region that is inhabited by a single representative household and has a single representative industry that uses one input factor, labour, which is owned by the representative household and is supplied inelastically, to produce one representative consumption good that is consumed by the household. Depending on the kind of analysis that one wants to undertake, it is necessary to relax some of these restrictive assumptions; for example, to use such a model for a meaningful analysis of tax policy, it needs to be extended in at least two, but preferably four directions.

First, the model needs to have more

than one factor of production, because the effect of taxes on the supply of different factors depends on each factor's elasticity of supply. Many AGE models employ at least two factors of production, labour and capital. Labour supply is usually modelled as the result of the household's utility maximisation, in which the household decides how much labour to supply and how much leisure to consume. Because the labour supply decision depends on the return to labour, the wage, such a model can be used to analyse the effect of taxes on labour income.

The decision to combine all other factors of production into a single factor "capital" is problematic. This single factor is an aggregation of mobile capital (for example, machines), immobile capital (structures), land, and infrastructure; because each of these factors has a different elasticity of supply and therefore responds differently to taxation, such aggregation can severely distort the results of the model. The supply of mobile capital is very elastic, because it can be easily moved and because it generally depreciates very rapidly. The supply of immobile capital is rather inelastic for reductions, because structures cannot be moved and depreciate slowly, but it is elastic for increases in the stock of capital. The supply of land is perfectly inelastic, because land can neither be moved nor does it depreciate physically. Infrastructure describes public goods that benefit many households and firms simultaneously, and that are often tax-exempt. Models that aggregate factors with such different characteristics are not likely describe the economic implications of tax policies on production accurately. A general equilibrium model should therefore differentiate between at least the described five factors of production.

If the required data for such a model are not available, it is of course necessary to aggregate some of these factors. However, one should resist the temptation of aggregating land and capital. A model that aggregates land and capital will overestimate the economic effects of taxing such a hybrid if it assumes that the hybrid is supplied elastically, and underestimate the effects if it assumes that the hybrid is supplied inelastically. In addition, such a model would be likely to ignore the fact that land will be used more efficiently when it is taxed more heavily.[10]

Second, the model needs to be dynamic (that is, describe multiple time periods simultaneously), because static (one-period) models are unable to capture the intertemporal distortions that are caused by taxing capital. New capital is financed through consumers' savings, and capital will depreciate if it is not replaced continuously. Static models require strong assumptions about the consumer's motivation to save and to invest, and economists consider it more appropriate to model the consumer's saving decision as a solution to an intertemporal utility maximisation problem.

Most AGE models assume that the total stock of capital is fixed within any given time period. Because capital depreciates over time and needs to be replaced, a model that describes only a single time period is likely to underestimate the long-term effects of taxes that provide disincentives to save and to invest.

Sufficient factor differentiation and the ability to describe multiple periods are the most important properties of AGE models that are used for the analysis of the economic effects of taxation. The following two additional model extensions are desirable, because they are likely to enhance the accuracy of the model's prediction. However, because they require additional data, their implementation may not always be feasible.

First, the model should have more than one representative consumer and more than one representative industry. The response of consumer behaviour to taxation depends on various socioeconomic factors such as household composition, wealth, age, and education, and models that aggregate consumers from very different socioeconomic groups into a single "representative consumer" face the difficulty of describing the exact behaviour of such an artificial representative unit.[11] Similarly, because different industries use different

combinations of input factors (for example, the service industry is very labour intensive, while the transportation industry is very capital intensive), their responses to the taxation of these factors will differ, and a more disaggregate model is more likely to describe the economic effects of such taxes accurately.

Second, if it is important to analyse the effects of regional taxes, or if the economy to be analysed consists of heterogeneous regions with different stocks of input factors, the model should describe these regions separately rather than aggregating them into a single (federal) unit. Regions with high stocks of capital are likely to have higher wages than regions with low stocks of capital, because the availability of more capital makes workers more productive. Depending on their location, workers are therefore likely to have different incomes, and to react differently to the taxation of labour income. Similarly, regions with large cities and large populations can be expected to have high demand for mobile capital, so that capital is moving into these regions and capital investment is high in these regions. The taxation of the return to capital in these regions is likely to alter the flows of capital as capital moves into less productive regions with lower taxes. Aggregating all regions into a single unit means that such region-specific responses are ignored, and such a model will underestimate the excess burden of taxes on labour and capital that is caused by inefficient migration of people and capital.

FOR THE ANALYSIS of alternative tax programs in Russia, Nicolaus Tideman's (2001) model is a very reasonable beginning. His model is dynamic (it covers 100 time periods), labour supply is determined as a solution to the consumer's utility maximisation problem, the model differentiates between capital and land, and it incorporates the fact that a tax on land value provides an incentive to use land more efficiently. His model permits analyses of several alternative tax programs. However, mainly because of the difficulties of collecting more detailed data for the Russian economy, his model has only a representative consumer and a representative industry that produces a single good, and it describes a single region. To the extent that Russia's consumers, industries, and regions are heterogeneous, the model is likely to ignore tax effects that arise from these differences. The larger these differences are, the less precise will be the analyses that are undertaken with his model. Because his, already quite detailed, model is programmed on an Excel spreadsheet, it could be very difficult to incorporate extensions that require additional orders of complexity. If the required data can be collected, a more promising approach would be to develop a model of the Russian economy in a more flexible programming language,[12] with multiple regions, sectors, types of capital, and types of consumers. Nicolaus Tideman and I have undertaken preliminary work that leads us to conclude that such a model is feasible,[13] but we have not yet produced a working model of a real economy.

Bibliography

Harberger, Arnold C. "The Incidence of the Corporate Income Tax." *Journal of Political Economy* 70: 215-240 (1962).

Henderson, J. Vernon. "Community Choice of Revenue Instruments." *Regional Science and Urban Economics* 24:159-183 (1994).

Hoyt, William H. "Competitive Jurisdictions, Congestion, and the Henry George Theorem: When Should Property be Taxed Instead of Land?" *Regional Science and Urban Economics* 21:351-370 (1991).

Ginsburgh, Victor and Michiel Keyzer. *The Structure of Applied General Equilibrium Models*, Cambridge: MIT Press (1997).

Pigou, Arthur C. *The Economics of Welfare*, London: Macmillan (1932).

Plassmann, Florenz and T. Nicolaus Tideman. "A Dynamic Regional Applied General Equilibrium Model with Five Factors of Production." *State University of New York at Binghamton Working Paper* WP9907 (1999).

Quesnay, François. "The 'General Maxims for the

Economic Government of an Agricultural Kingdom," in Ronald L. Meek, pp.231-262 (1963) [first published in 1756].

Shoven, John B. and John Whalley. *Applying General Equilibrium*, Cambridge: Cambridge University Press (1992).

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*, Indianapolis: Liberty Fund (1981) [first published in 1776].

Tideman, T. Nicolaus. "Analysis of Alternative Tax Reforms for Russia." *Mimeo* (2001).

References

1 Quesnay (1963 [1756]), p.232.

2 Smith (1981 [1776]), p. 844.

3 Such a tax is known as a Pigouvian tax (see Pigou, 1932).

4 See, for example, Harberger (1962).

5 It is well known in the local public finance literature that a poll tax can be implemented by local communities to distort people's migration decisions (Hoyt, 1991; Henderson, 1994).

6 If the government's demand for revenue exceeds the amount that can be obtained through these taxes, then it is necessary to impose additional taxes that have excess burdens.

7 Applied general equilibrium models are generally developed within the neoclassical framework. On the consumer side, agents are assumed to be rational and utility maximising; they are endowed with sets of factors that they exchange, either among each other or with producers, to maximise their utility. Producers transform agents factor endowments into output, which they sell to obtain funds to buy factors from agents. Markets are generally assumed to be competitive, so that producers do not make positive profits after paying all factors their market returns, and very often production is assumed to take place under constant returns to scale.

8 For recent summaries of the literature on Applied General Equilibrium models see Shoven and Whalley (1992) and Ginsburgh and Keyzer (1997).

9 "Calibration" is the process of choosing values for the constants of the model so that the model's equilibrium levels of output, employment etc. match those of the economy that it is meant to describe.

10 Because it is costly to remove stationary capital from land, land speculators have an incentive to leave land idle in the hope of selling it for a higher profit in the future. A tax on land value increases the cost of leaving land idle, and thereby leads to reduced land speculation and more efficient use of land.

11 Theoretically, a highly aggregate model is able to describe the aggregate effect correctly if it determines the weights of individual groups within the aggregate unit according to the group-specific response to the particular policy that is to be analysed. However, not only would such a model require different aggregations for the analyses of different policies, but it would also be necessary to know the correct response of each individual group to any policy of interest; such detailed information is unlikely to be available.

12 Such a model could be developed in a programming language such as C or Fortran, or possibly within the user-interface GAMS, using extensions of traditional AGE methods. C and Fortran are (low-level) programming languages that are extremely flexible and that can be used to develop any software. Their greatest advantage is their enormous flexibility, their only disadvantage is that they are (relatively) difficult to use. GAMS (General Algebraic Modelling System) is a (high-level) programming language for the solution of mathematical problems. It is often used to develop AGE models, because it has a large number of pre-programmed routines that are useful for this purpose. GAMS is easier to use than C or Fortran, but because it is more specialised than low-level languages, its possibilities are more limited, and it may not be possible to use GAMS to program the type of model discussed here.

13 See Plassmann and Tideman (1999).