

APPENDIX V: METHODOLOGY

The tax incidence results reported in this study for calendar year 1995 were obtained from the Institute on Taxation and Economic Policy (ITEP) microsimulation tax model. Three major components of the model—relating to state income taxes, sales and excise taxes and property taxes—are fully integrated in order to allow for a comprehensive study of the combined incidence of the major taxes affecting individuals and businesses at the state and local level. This appendix describes the components of the ITEP model and how they are linked together; the various data sources used; calibration and extrapolation of the baseline data; and assumptions relating to the incidence of specific taxes.

The ITEP Personal Income Tax Model

Overview

Data from federal income tax returns form the primary basis of the ITEP income tax model. Approximately 400,000 individual tax records provide much of the necessary income, deduction and demographic data. This data was obtained from the Statistics of Income (SOI) Level III public use data file—a stratified random sample of federal tax filers chosen in such a way as to provide statistically valid estimates of state-by-state income, deductions and taxes. To this population of federal tax filers was added additional information on the non-filer population from the Decennial Census to create a comprehensive database of approximately 700,000 individual records calibrated to match state-by-state population, income and other aggregate totals for calendar years 1988 to 1998. The ITEP microsimulation model is one of the largest of its kind in the world.

Statistical Matching of Census Records

The first step in the construction of the database was to statistically “match” relevant data from the Census to individual income tax return data on a state-by-state basis. This was necessary for at least two reasons. First, many items relevant to the proper calculation of state tax liability are not available on federal income tax returns. For example, some states allow special deductions and/or exemptions based on the age of the taxpayer and this information is largely absent on the

federal return (it is possible, however, in most cases, to identify those taxpayers who are over 65). Additionally, information on income that is not reported on the federal return is needed for calculation of total income for classification purposes.

Secondly, federal tax return data only report information on those who file federal tax returns. Since this population is different than the population subject to all state and local taxes, the data need to be supplemented to account for the (federal) non-filing population.

Statistical matching of information from tax and non-tax sources is a well-established method for augmenting available data as well as extending the range of analyses that can be performed on issues affecting policy. While there are several different methods currently in use that perform this task, most share a number of common characteristics which were followed in the construction of the ITEP model.

The following procedures were implemented separately for each of the fifty states and the District of Columbia. First, hypothetical “tax units” were constructed on the census database (“the augmentation file”), based on incomes, family composition, living arrangements and other relevant demographic factors. Next, these tax units were classified as being a federal tax filer or non-filer based on the income and demographic information available. Special care was taken so that all relevant household

information was utilized. For example, for a household with more than one tax filing unit it was assumed that aged and dependent care exemptions were allocated in the most likely way.

In the next step, the census file is “aligned” with the Level III SOI file (“the base file”) in the following manner: tax units in each file are categorized along several dimensions according to variables common to both files. These variables included age, income, family size, home ownership, number of earners and level of earnings for each and the presence of certain types of income. Once the relevant categories were assigned, records were matched while preserving the essential income and demographic make-up of each file. This was accomplished by adjusting sampling rates—a technique known as “constrained merging.”

1995 State Personal Income Tax Law

Forty-three states and the District of Columbia impose some type of personal income tax. Since a majority of these states are linked (“coupled”) to the federal income tax in some way, federal tax data are a natural starting point for most types of policy analyses. For each state with a personal income tax, the ITEP model computes, when applicable, state adjusted gross income, itemized deductions, personal exemption amounts, dependent exemption amounts, taxable income, tax credits and final tax liability for each return in the sample, based on the tax law in effect for each state for calendar year 1995.

For states that allow two-earner married couples to file separately on the same return, a few rules of thumb were adhered to: (i) this option was assumed to be chosen if it resulted in a reduced tax liability; (ii) wage and salary income was apportioned accord-

ing to the information matched from the census records; (iii) other income amounts were allocated in such a way as to minimize tax liability to the extent allowed by state law; (iv) deductions were allocated in order to minimize tax liability to the extent allowed by state law; (v) personal and dependent exemptions were allocated to minimize tax liability to the extent permitted by law.

Extrapolation and Calibration

Forecasts of the number of returns, population, income, deductions, state tax liability and federal tax liability were constructed for all fifty states and the District of Columbia based on historical data published by the Internal Revenue Service (through 1993), national macroeconomic data from the Commerce Department, national macroeconomic forecasts published by the Congressional Budget Office (CBO) and information supplied to ITEP by various state revenue departments. Since 1993 is the last year data were available relating to the distribution of selected income and deduction items by state, the extrapolations were constrained to hit these totals.

The extrapolation was accomplished in two stages. In the first stage, all income and deduction items were adjusted by per capita growth factors, first with respect to national totals and next with respect to state-specific factors. In the second stage of the extrapolation, population weights were adjusted to hit aggregate control totals for population, returns and selected items of income and deductions.

As a final check, for states with an income tax, tax liability estimates from the ITEP model were compared with the latest estimates from the states. (In most cases, fiscal year 1995 was the most recent year available).

The ITEP Sales Tax Model

Overview

The ITEP Sales Tax Model calculates the distribution on in-state residents in each state of state and local taxes imposed on the sale or use of specific commodities or services. These taxes include general sales taxes (including those imposed at the local level), excise taxes and numerous special taxes affecting specific industries (hotel taxes, for example).

The model computes the total amount of these taxes for three separate groups of taxpayers: residents of a particular state; businesses located in the state; and visitors to the state. This breakdown is necessary since it is likely that the overall burden of a particular tax on in-state residents depends on how much of the revenue is borne by each of these three groups.

Taxes Paid by Residents

Data Sources

Information on the detailed consumption patterns of U.S. families is available from the Consumer Expenditure Survey, U.S. Department of Labor, Bureau of Labor Statistics (CEX). The CEX is a quarterly survey, conducted on a rotating basis, of approximately 7,800 households per quarter. Each household is attempted to be interviewed for four consecutive quarters. In the ITEP model, annual consumption expenditures were constructed for each of the families interviewed from the fourth quarter of 1991 through the first quarter of 1993. To facilitate spending to income comparisons, only those families that responded to all four interviews were included in the analysis. Additional information on the income and demographic make-up of respondents was also included.

Methodology

Detailed consumption amounts were imputed to each household in the database by estimating the relationship between

various consumption components and the characteristics of each household using standard econometric techniques. Regressions were run on the CEX data after first: (i) eliminating those households for which complete income information was unavailable; (ii) correcting for the “top-coding” of high-income returns; and (iii) discarding a small number of outliers. The analysis was done separately on homeowners and renters.

The imputation process proceeded in several stages. First, expenditures on “big ticket” items such as automobiles and consumer durables were estimated separately since these items, while an important part of each household’s expenditure budget, represent a transitory component of overall annual consumption. Since most families do not purchase these items every year, special econometric techniques were used to obtain consistent parameter estimates. Once these estimates were obtained, the probability that a family actually made such a purchase was computed and compared with a uniformly distributed (pseudo-)random number to determine if an imputation was to be made to this record. The size of the purchase was then calculated according to the estimated parameters.

Next, estimates of total consumption less these major purchases were imputed for every family in the database using the results from an ordinary least squares (OLS) regression, relating this consumption amount to a vector of independent variables such as family size, income and the ages of the principle earners in the family. Since this imputed amount represented the mean consumption for a family with these same characteristics, a random error term was added to each amount to impart an additional degree of variance to the final imputed value. The error term was drawn from a normal distribution with a zero mean and a standard deviation equal to the standard error of the estimated regression equation (separately for homeowners and renters).

Once estimates of total non-durable consumption were made, this amount was allocated across major consumption categories: food, housing, clothing, transportation and other consumption. This was accomplished by estimating the share of each of these items as a fraction of each household's spending and using the resulting parameter estimates to impute budget shares to each family unit (dependent filers were excluded from this procedure). Since the sum of these shares must necessarily sum to 1.0, one equation is redundant and was dropped from the estimation process. This system of equations was then estimated by a "seemingly unrelated regression" (SUR) approach and the resulting variance-covariance matrix used to generate a vector of multinormal random numbers, which when added to each share equation yielded the final set of consumption shares for each tax return.

More detailed consumption categories were imputed to taxpayers in a similar manner as that outlined above: share equations were estimated from the CEX in order to further subdivide each consumption category (food was further subdivided into food consumed at home and food consumed away from home, for example). The estimated coefficients were used to impute average shares given particular household characteristics and a vector of multinormal random error terms was added. In total, imputations for 61 different categories of consumption were imputed to each household in this manner. In certain cases where items subject to the sales tax were at a finer level of detail than could reasonably be estimated in the above manner, the appropriate category was further subdivided by spending shares contained in the CEX.

Extrapolation and Calibration

Since the ITEP microsimulation model ages the income, tax and demographic data in order to track macroeconomic aggregates over the simulation period, imputed consumption amounts will also reflect this aging process but only to the extent they are

captured by the list of predictor variables. In order to account for other unexplained components, a simple multiplicative adjustment factor was applied to total consumption in order to more closely track estimates of personal consumption expenditures (PCE) contained in the National Income and Product Accounts (NIPA). Where accurate control totals were available, the model was targeted to hit these NIPA aggregates. No attempt, however, was made to adjust the individual components for those categories of consumption subject to significant differences in definition and reporting.

Since certain items of personal consumption exhibit regional variation that are not captured in the CEX, these items (fuel oil in the Northeast, for example) were adjusted via state-by-state factors obtained from auxiliary sources to arrive at final estimates of total personal consumption.

Taxes Paid by Business

Data Sources

Businesses are taxed in most states having a sales tax on the purchase of goods and materials used in the production of final output as well as certain capital expenditures. Most states, however, exempt raw materials that are a part of the final product from the sales tax base. In order to accurately compute the sales taxes attributable to business, it was necessary to know the industry composition each state; the intensity of each industry within the state; and the inputs into the production process. ITEP's business sales tax model accomplishes this by utilizing disaggregated industry-specific estimates of Gross State Product (GSP) prepared by the U.S. Department of Commerce in conjunction with the Benchmark Input-Output (I-O) tables of the United States.

Methodology

For each state, a regional I-O model is constructed from the Benchmark "Use" table by assuming the technical coefficients remain unchanged across states and

adjusting the inter-industry flows by the “location quotient” method: for each industry, the entries in the national I-O table are adjusted by a factor, LQ_i , representing the fraction of that industry’s contribution to GSP in the state to national totals for the industry. Estimates of final demand were obtained in a similar manner. The commodity composition of intermediate production was mapped into PCE consumption categories using the “consumption bridge” in the I-O accounts.

Numerous exemptions from the sales tax base are provided to businesses in most states. To account for these exemptions in the calculation of sales taxes paid, business inputs were further subdivided into five mutually exclusive categories: raw materials used in the final product, utilities, machinery, inputs used in agriculture and all other inputs. These exemptions comprise most of the sales tax exemptions currently in effect at the state level. Detailed purchases of commodities were obtained in this manner for each of 49 industries operating in each state.

Extrapolation and Calibration

The Benchmark I-O table was constructed for 1987. In order to get estimates of industry sales and purchases for out-years, forecasts of state GSP by industry were used to age the regional I-O models. These forecasts were obtained from the U.S. Commerce Department.

Taxes Paid by Visitors

State-by-state estimates of expenditures by travelers are provided by the U.S. Travel Data Center. These expenditures are broken down into six categories: public transportation, auto transportation, lodging, food service, entertainment and recreation, and general retail sales. In order to properly reflect each state’s sales tax base, these categories were further subdivided based on proportions contained in the CEX with minor adjustments reflecting the likely composition of expenditures by tourists. The last year data are available was 1993; expenditure data for subsequent years were adjusted to reflect the change in the consumer price index (CPI).

The ITEP Property Tax Model

Overview

ITEP’s property tax model is unique in that information from individual tax returns, including data relating to property tax deductions, is coupled with Census data relating to home values so as to provide an integrated tool for policy analysis. This aspect of the model is particularly important for the analysis of state tax policy incidence since many states attempt to mitigate the effects of high property taxes via credits and/or exemptions which are income based (such as “circuit breakers”).

Home Property Taxes

For tax returns where itemized deductions are taken, the starting point for the analysis is itemized deductions for real property taxes. This property tax value is

used in conjunction with the home value obtained from the Census match and state-specific information on rates, assessment practices, homestead exemptions and other tax-reduction provisions to calculate all of the components of a property tax calculation for each record including: tax rate, an assessment ratio and, where applicable, a homestead exemption. With this information, tax is then recalculated for 1995 incorporating extrapolations of value and incorporating tax law changes occurring since the base year of the data.

The same approach is used for non-itemizers except that the property tax amount used is from the Census. All of the results, particularly assessed values, homestead exemptions, taxable value and property taxes paid are compared to aggregate data from the state, where available.

Residential Rental Property Taxes

Rent paid (from the Census) is multiplied by a factor to calculate the value of rental units. Average state tax assessment ratios and rates for residential rental property are multiplied by this value to calculate the tax. Half of the tax is treated as paid by renters. The other half is treated as paid by owners (see taxation of business, below).

Ad-Valorem Automobile Taxes

For itemizers, the personal property tax deduction is used to impute auto values. With this information, tax is then recalculated for 1995 incorporating value extrapolations and tax law changes since the

data year. For non-itemizers auto value is imputed using income and number of household members. Aggregate results are compared with data from the state, where available.

Intangible Property Taxes

For itemizers, personal property tax deductions are used to impute the value of intangible assets. With this information, tax is then recalculated for 1995 incorporating value extrapolations and tax law changes since the data year. For non-itemizers the value of intangibles is imputed using income from intangibles. Aggregate results are compared with data from the state, where available.

Business Tax Incidence

Overview

The ITEP Model relies on a combination of microsimulation, imputation and calculation to allocate taxes. In the case of taxes initially paid by business, the primary allocation tool is the last of the three.

Corporate income taxes, other corporate capital taxes and business property taxes

The total amount of corporate income taxes and corporate taxes based on capital values for each state for calendar 1995 were calculated based on the most recent available revenue data from each state. The figures include local corporate income taxes where applicable. Business property taxes for each state for calendar 1995 were computed as a residual amount based on our estimates of total property tax payments for the year for each state, less the amount paid by homeowners (as calculated by the property tax microsimulation model).

Allocation rules

Property taxes on residential rental property: Half of total business property taxes on residential rental property were assigned to tenants. This portion of the business property tax was distributed based on rents paid, and is included in the "Property taxes, individuals" section of the report. The remaining half of property taxes on residential rental property was distributed by capital ownership, as described below.

General rule for other business income, capital and property taxes: The remaining corporate income, capital and business property taxes were generally treated as taxes on capital, and assigned to capital income, in and out of state according to the following rules:

Exporting: In assigning taxes to capital income, each state retained its share of national capital income plus an adjustment to account for the greater likelihood of residents owning in-state taxable capital. For

residential rental property taxes, this adjustment averaged 50% of the remaining portion of the tax. For other business property taxes, the adjustment averaged 20%.

“Excessive” capital taxes: In states that impose unusually high taxes on capital, those taxes may not be entirely borne by capital owners. Instead, they may be shifted back to wages or forward to consumption. We computed the total amount of each state’s corporate income and capital taxes and non-residential business property taxes as a share of output by industry type. In cases where one of these computed taxes was significantly above the national median, we assigned the excess to either in-state wages or in- and out-of-state consumption, depending on the type of activity.

Sales & excise taxes initially paid by businesses

Estimates of the dollar amount of sales and excise taxes paid directly by businesses were computed according to the ITEP Business Sales Tax model (described above) for taxable purchases of business inputs. In general, sales & excise taxes initially paid by businesses were assigned to consumers based on shares of total consumption, with the following special rules.

Exporting: Sales and excises taxes initially paid by businesses were divided into those

taxes paid by businesses that produce output sold principally in national versus those sold principally in domestic markets. This distinction was made based on the following factors: (1) the particular industry producing the product; (2) the share of each industry’s contribution to each state’s Gross State Product; and (3) special considerations specific to a particular industry within a state. Taxes on domestic market items were kept in-state (except for amounts paid by visitors). Taxes on national market items were assigned to national consumption, with an adjustment retaining, on average, 15% of the tax in-state (on top of the state’s share of national consumption).

“Excessive” sales and excise taxes on national market items: In states that impose unusually high sales and excise taxes on national market activities (in relation to the national median), those taxes may not be entirely borne by consumption because of competitive factors. Instead, they may be shifted backward to in-state wages or capital. Total national market business sales and excise taxes were computed as a share of each state’s national market GSP. In cases where these national market business taxes were significantly above the national median, we assigned the excess half to in-state wages and half to capital. The latter amounts were distributed according to our rules for allocating capital taxes.

The Study Population

In this study we report taxes paid by non-elderly married couples. The reasons for reporting on a subset of the population are:
Income classification across different types of family units can prevent accurate evaluation of tax equity. For instance, a single person with income of \$20,000 has a far better ability to pay taxes than does a married couple earning \$20,000. Thus, the equity of, say, an 8% burden on an income group with incomes ranging from \$18,000 to \$22,000 is easier to judge if only a subset of the population is included.

Income classification across different types of families can distort the patterns within subsets. An overall distribution of income is dominated by single individuals in the lower income ranges and dominated by married couples in the upper income ranges. Thus, an assessment of taxes by income done in a combined table could measure distinctions in the treatment of single individuals versus married couples instead of differences by income groups. Although this may be an issue worthy of study, it is not the purpose of this report.

Elderly people's income often does not reflect their wealth. Although relative to other elderly, income may be a good measure of relative wealth, a portion of the elderly population that would appear low- or middle-income in an overall distribution is, in fact, quite well-off. Putting the elderly in a combined table can thus distort the overall results.

The reason we chose non-elderly married couples is that it is a subset of the population that pays most of the taxes, that represents a majority of the population (approximately 55% of the population lives in

such households) and which a very high portion of the population is a part of for a significant portion of their lives. That is not to say that tax incidence of other populations would not be of great value, and we hope to report on such populations in subsequent studies.

Finally, it is important to note that studies of the populations excluded from this study would, based on preliminary analyses we have done, and the work of others, show quite similar patterns of tax incidence to those we have shown here.

Comparisons With Previous Citizens for Tax Justice Studies

The methodology used in this study is very different than that used in previous Citizens for Tax Justice analyses of state and local taxes. Although the conclusions and patterns presented here are consistent with previous findings, direct comparisons of specific data for the purpose of indentifying trends will not produce accurate results.