A Online Appendix of:

Distributional Tax Analysis in Theory and Practice: Harberger Meets Diamond-Mirrlees

by Emmanuel Saez and Gabriel Zucman

A.1 Practical Considerations For Current-Tax Analysis

This appendix provides a tax-by-tax discussion of the practical implementation of distributional current-tax analysis for cases that are not immediately covered by the general principles outlined in Section 2.

Taxes on intermediate goods. Some consumption taxes (such as tariffs, taxes on alcohol and fossil fuels, and business turnover taxes) are levied on intermediate rather than final goods. Intermediate goods taxes are small, less than 3% of total tax revenue in the United States. Most countries have replaced turnover taxes by the value-added tax which only taxes final consumption.[38] Because taxes on intermediate goods distort production prices, there is no direct model guidance on how to assign these taxes for distributional current-tax analysis.

In our view, the best way to proceed is to treat these taxes as consumption taxes on the final goods eventually produced using the taxed intermediate goods. For example, a tax on wholesale beer will be assigned to final beer consumers (as part of the post-tax beer price), a tax on jet fuel to the consumers of airplane travel. A more complex case involves turnover taxes on natural resource extraction that many extracting countries impose. If the marginal cost of extraction is equal to the selling price (no pure profits), the tax is akin to an intermediate goods tax. However, if the marginal cost of extraction is lower than the selling price (e.g., oil extraction in Saudi Arabia, where marginal costs are much lower than the global oil price determined by the marginal producer), royalties are akin to a tax on the pure profits of extracting companies. Practically, one needs to assess whether the royalty is assessed on a resource for which production is closer to the no pure profit vs. pure profit benchmark. In the case of US oil and gas extraction, marginal costs are significant and we treat royalties levied by US governments (0.2% of government revenue) like taxes on other intermediate goods.[39]

Taxes on depreciable capital assets. Assets used in production are subject to property taxes. If the asset does not depreciate (e.g., land) the tax is fully assigned to the ultimate owner

[38] Intermediate goods taxes create production inefficiencies and the Diamond and Mirrlees (1971) model shows that they should not be used.

[39] In 2021 taxes on the extraction of natural resources such as oil and natural gas, called severance taxes, generated $13.5 billion in revenue (NIPA Table 3.5), out of $6.3 trillion in government tax revenue.

of the asset. If the asset depreciates (e.g., a building) then the depreciating part of the asset is like an intermediate good: it is consumed during the production process. The corresponding tax is allocated like other taxes on intermediate goods, i.e., to consumers of the corresponding final goods. For example, if Amazon uses up 1/40 of its warehouses each year (straight-line depreciation over 40 years), then 1/40 of the annual property tax paid on these warehouses are included in the consumption taxes on Amazon products sold to final consumers. In practice, because the bulk of taxes on capital assets are property taxes on buildings and land, which have long or infinite lives, business property taxes can be fully assigned to business owners.[40]

Carbon taxes. Carbon taxes may become important during the transition to clean energy. Because both consumption and investment decisions are responsible for carbon emissions, a general carbon tax covering all forms of emissions should be allocated to both consumers and owners, in proportion to carbon emitted. In the case of emissions due to investment (e.g., a warehouse built with cement), the intermediate-goods logic described above continues to apply. Since assets are partly consumed during the production process, part of the tax should be allocated to the consumers of the final goods produced with the depreciating assets. Overall, our framework assigns carbon taxes to consumers in proportion to consumption of final goods and consumption of fixed capital, and to business owners in proportion to net investment (i.e., gross investment minus consumption of fixed capital).

Because business ownership is more concentrated than consumption, with this methodology carbon taxes are more progressive than when the assignment is only based on the consumption of final goods (ignoring investment), the conventional approach (see Carloni and Dinan, 2021, for a survey). But carbon taxes are less progressive than in the methodology of Chancel (2022) and Chancel and Rehm (2023), where carbon emissions are allocated to consumers for consumption goods and business owners for *gross* investment (instead of *net* investment as we propose). In the United States, net domestic investment is only about 25% of gross domestic investment. Our method is thus approximately 1/4 of the way between the conventional method and the Chancel method.

Inheritance, gift, and estate taxes. Taxes assessed on the transfer of wealth can hurt the welfare of two parties: the donor and the donee.[41] They could be assigned to either. We follow the conventional approach that assigns taxes to donors. This can be rationalized by the fact that

[40] In the case of residential property taxes, for owner-occupiers the owners and the consumers are the same individuals, so there is no assignment issue. For rented housing, the part of the property tax corresponding to the annual depreciation of the structure should conceptually be allocated to the consumers of housing services (the renters). This part, however, is very small (1.25% of the property tax assuming (i) straight-line depreciation of the structure over 40 years and (ii) that land, which does not depreciate, represents half of the taxable value of the house) and can be neglected in practice.

[41] Piketty and Saez (2013) propose an optimal inheritance tax model where both welfare effects play a role. In a dynastic model of Barro-Becker, donor and donee are part of the same dynasty, but in the real world individuals

the potential negative impact on donors is the one that usually raises most concerns (transfer taxes harm the property rights and incentives of donors to accumulate wealth), while the cost for donees is secondary (as they benefit from a transfer through no effort of their own).[42]

Transaction taxes. Some countries impose taxes on specific transactions such as real estate transactions, or financial transactions. The simplest treatment is to allocate such taxes to the buyer side of the transaction (and make it flow to the ultimate individual owner if an intermediary such as a business is buying the asset). This naturally extends our treatment of consumption taxes where consumption taxes charged on second-hand goods are also assigned to the buyer.[43] If turnover is fast (as is often the case with financial transactions), allocating to buyers vs. sellers does not make much of a difference.

Progressive consumption taxes. A progressive consumption tax that exempts net savings from taxation and adds net dissaving to the tax base (i.e., that extends the traditional pension treatment to all forms of savings) is allocated to individuals based on their consumption. As savings are concentrated at the top of the income distribution (Saez and Zucman, 2016) with negative savings at the bottom and positive and large savings rate at the top, moving to a progressive consumption tax would be regressive when distributional impacts are assessed relative to income percentile.[44]

Flat taxes. Flat taxes have been proposed in the US tax debate by Bradford 1986 (the X-tax) and Hall and Rabushka 1985 (the flat tax). This "flat tax" is a tax on wage income combined with a cash flow tax on business profits with no deduction for interest income payments and full expensing of investment instead of depreciation of capital assets over their lifetime as in regular corporate taxes.[45] Using our methodology, the flat tax would be assigned on the corresponding wage earners and the corresponding business owners.

While the "flat tax" is economically equivalent to a flat consumption tax such as a VAT from a dynamic perspective (and ignoring the exemption for low earners built in the flat tax),

matter separately from dynasties (and indeed to the best of our knowledge, no distributional tax table has ever been presented for dynasties).

[42] Arguments in favor of assigning taxes on donees can also be provided. For example, if bequests are accidental, then donors do not care about transfer taxes and only donees are affected.

[43] The convention in national accounts is that if a second-hand good is resold through a business, it is seen as a business activity with the used good being an input and the resold used good being like a new good with the difference in prices reflecting value added: the cost of buying and reselling the used good for the business, and the value of reallocating the good to a consumer with higher value on the consumer side.

[44] Proponents of consumption taxation might argue that individuals should be ranked by consumption rather than income when assessing progressivity. To our knowledge, such distributional tables based on consumption have not been produced, in large part because there is no good micro-data in the United States measuring both income and consumption especially at the top of the distribution.

[45] TCJA provides full expensing for five years 2018-2022 with a phased-in return to depreciation over 2023-2027.

the distributional impact is quite different when measured on an annual basis. A worker who saves most of his income consumes little and hence pays no consumption tax, but would pay the "flat tax" on wage earnings. As highly paid workers save more than low paid workers, the flat tax will be more progressive than the VAT on an annual basis.[46] The "flat tax" exempts investment while the consumption tax exempts savings. Investment is made by business owners who may be different from savers but both business owners and savers are concentrated toward the top of the distribution. Therefore, on net, the flat tax is likely to be more progressive than the VAT, measured on an annual basis. Naturally, from a dynamic perspective, the two taxes generate the same budget sets and hence are formally equivalent (See Auerbach 2019 for a recent exposition). However, if households face borrowing constraints or do not plan according to the classic intertemporal utility model, this equivalence is lost.

Taxes on mixed business income. Business income is a mix of labor income (the labor effort of the owner) and capital income (the return on the business assets). Neither national accounts nor income tax data can separate cleanly the two components. How then should we assign the corporate income tax on a closely held business or the individual income tax on pass-through businesses? With our methodology, such taxes are assigned directly to the owners themselves who supply both the labor and the capital so we do not need to separate labor and capital to assign taxes either. CBO assigns 25% of the corporate tax to workers but it assigns 100% of the tax paid by a passthrough business to its owners (because it allocates individual income taxes to each taxpayer individually). Hence, a pure change in organizational form with no change in economic activity, such as a change from a sole proprietorship to a C-corporation, increases the tax rate on workers nationally, which is not satisfactory conceptually.

Foreign taxes. One limitation of both the conventional approach and ours is that cross-border corporate income tax payments are ignored. Only corporate taxes collected by the US government are allocated to individuals. This is because government agencies are interested in distributing US federal tax revenues, while the distributional national accounts literature is interested in distributing national income—and foreign corporate taxes are not part of national income. In reality US individuals pay corporate taxes to foreign governments, and some of the corporate taxes collected by the United States are paid by foreigners. In recent years the two flows broadly offset each other and our series thus capture the effective rate paid by US individuals globally. However this was not the case historically. It would be valuable to develop distributional current-tax series adding back net cross-border corporate income tax payments,

[46]Similarly, for private pension arrangements in the US, a Roth IRA is equivalent to a traditional IRA from a lifetime perspective. But on an annual perspective, if savers who get the tax exemption through the traditional IRA have higher incomes than retirees who get the exemption through the Roth IRA, the traditional IRA is less progressive than the Roth IRA. Viard and Carroll (2012) note that the flat tax is like a Roth IRA while the VAT is like a traditional IRA.

a task we leave to future research (Zucman, 2023).

How to rank individuals? Traditional distributional tables ranks individuals (or families) by annual pre-tax income. This is justified if annual pre-tax income is indeed the best measure of economic status. Other rankings are conceivable such as changing the time frame (such a month, multi-years, or even a lifetime) or changing the variable to after-tax income, consumption, or wealth. There is no definitive or right answer to this question. Different measures might work best for different purposes. At the high end, wealth plays a role over and above income to measure economic status. A CEO earning $50 million/year with no accumulated wealth is not in the same economic class than a wealthy owner making $50 million/year out of a fortune of $1 billion. This would call for factoring wealth over and above the capital income it generates in some way. Consumption becomes an almost irrelevant variable at the very top as even lavish personal consumption is going to be small relative to wealth for billionaires or deca-billionaires. At the low-end, transfers play a large role so that after-tax and transfer disposable income is likely to be a more meaningful measure of economic well-being than pre-tax income.[47] Even at the low end, consumption may not be a better measure of economic well-being than disposable income (available for consumption and savings) as the ability to save is clearly a marker of economic security and hence well-being. If our view, economists have spent too little time thinking through these important and non-trivial issues.

A.2 Current-Tax Analysis in Distributional National Accounts

This Appendix provides guidelines for the application of distributional current-tax analysis in the context of distributional national accounts, economic statistics that allocate all national income, taxes, and transfers to individuals. Section 2 and Appendix A.1 provide general principles and tax-by-tax discussions. Here we focus on the subtle issue of how to deal with indirect taxes for the measurement of inequality and for the estimation of effective tax rates by income groups. We discuss both issues in turn.

Pre-tax and post-tax incomes. National income includes indirect taxes. To estimate the distribution of national income, the most sensible approach is to first estimate the distribution of national income excluding consumption taxes (i.e., basic-price national income), and then gross up income levels proportionally (i.e., with no impact on the distribution of income). What follows details the reasoning.

At the micro-level, pre-tax income y and post-tax income c are related as follows:

$$c + t_c = y - t_y + g, \tag{5}$$

[47] Pre-tax and after-tax income rank might differ substantially if transfers are targeted to specific groups.

where y is pre-tax income (from labor and capital), t_y taxes paid on labor and capital generating pre-tax income y, g are transfers from the government, c is consumption—exclusive of consumption taxes paid—plus saving and t_c are taxes paid on consumption. The relevant concepts for inequality analysis are y (pre-tax income) and c (post-tax income). Total income $c + t_c$ is less interesting because it is an intermediate concept that includes taxes on consumption.

Importantly, equation (5) can be defined using broad or narrow definitions of income, consumption, and government transfers. At the broadest level: y includes all pre-tax income from labor and capital (labor income cash or in-kind and capital income distributed or retained within a business); g—and hence c—includes all forms of public spending (including collective consumption expenditures such as defense, education, etc.).[48]

Taking the broadest definition of income, equation (5) can be aggregated across individuals. Using capital letters for the macro level, we have:

$$\text{National Income } NI = C + T_c = Y - T_y + G. \tag{6}$$

c and y aggregate to $NI_f = C = Y$ *basic-price national income* (national income minus taxes on consumption) while $c + t_c$ aggregates to national income NI. $t_c + t_y$ aggregates to total taxes in national income $T_c + T_y$. As G includes all forms of government spending net of the government deficit, G also aggregates to total taxes $T_c + T_y$ in national income. As a result, $Y + T_c$ is also national income.[49]

As both y and c, the most relevant concepts for inequality, aggregate to basic-price national income, it is the most natural aggregate concept for distributional analysis. Basic-price national income measures income at pre-tax prices (i.e., prices before consumption taxes) while national income measures income at post-tax prices (prices inclusive of consumption taxes).[50]

Let us denote by $\tau_c = T_c/NI_f = T_c/C = T_c/Y$ the aggregate consumption tax rate so that $NI = (1 + \tau_c) \cdot NI_f = (1 + \tau_c) \cdot C = (1 + \tau_c) \cdot Y$. In general, $\tau_c \simeq 10 - 15\%$ in advanced economies. It is possible to blow up c and y uniformly by a factor $1 + \tau_c$ so as to aggregate to national income, which is more widely used in national accounting than basic-price national income, without affecting inequality indexes. The drawback is that this makes the incomes less

[48] As collective consumption expenditures provided by the government are hard to assign across individuals, it can also sometimes be useful to net them out on both sides of equation (5). Denoting them by cg, we have $g = g' + cg$ where g' are transfers from the government that can be individualized and valued individual by individual (such as cash and quasi-cash transfers) and $c = c' + cg$ where c' is disposable income of the individual so that $c' + t_c = y - t_y + g'$ which is a narrower definition of post-tax income.

[49] Indeed, national income is built as the sum of labor and capital income (which equals basic-price national income) plus all indirect taxes on products (i.e., what we assign as consumption taxes). Minor caveat: in national accounts, property taxes are counted in taxes on products while we think it is better to count them as part of capital income of owners (except for the depreciation piece, cf. Appendix A.1 above).

[50] In a closed economy, basic-price national income can buy national production at pre-tax prices (but not at post-tax prices). National income can buy national production at post-tax prices. This explains the unintuitive fact that consumption taxes have to be added to basic-price national income to get to national income even though individuals use their factor income to purchase goods and pay consumption taxes.

concrete relative to the incomes received by people.[51]

Tax rates and transfers. We now turn to the issue of how consumption taxes should be treated for the estimation of effective tax rates. In brief: consumption taxes should be allocated to consumers (as explained in the paper), but the portion of consumption taxes paid out of transfer income are best treated as reducing transfer income rather than as taxes.

To see this, note that in equation (5), t_y and t_c are the taxes paid by the individual on her income y and when consuming (or saving) disposable income $y - t_y + g$. Therefore, it makes sense to assign t_c separately to $y - t_y$ and g in proportion of the taxable consumption generated by each component.[52] Hence we split t_c into t_{cy} the consumption tax assigned to $y - t_y$ and t_{cg} the consumption tax assigned to g and re-write (5) as:

$$c = y - t_y - t_{cy} + g - t_{cg}. \tag{7}$$

The net transfer received is $g_n = g - t_{cg}$ and the total tax paid on pre-tax income is $t = t_y + t_{cy}$. This tax concept is the most natural one to estimate effective tax rates by income groups.[53] It avoids the issue of assigning very large tax rates to individuals at the bottom of the pre-tax income distribution with very low income y relative to transfers g and who pay consumption taxes on their consumption out of transfer income.[54] It makes sense to measure transfers as $g_n = g - t_{cg}$, i.e., net of consumption taxes paid.

At the macro level, in our view, the economy-wide tax rate is best defined as $(T_y + T_{cy})/NI_f$ i.e. as taxes T_y paid on factor income Y plus consumption taxes T_{cy} exclusive of consumption taxes paid on government transfers divided by factor income $Y = NI_f$. Note that this macro tax rate is generally not exactly the same as the ratio of total taxes paid divided by national income $(T_y + T_c)/NI$, a commonly used measure of the macro tax burden at the country level because the traditional measure weighs sectors based on after-tax prices while our proposed measure weighs sectors based on pre-tax production prices.[55]

[51] Furthermore, it is not possible to move from $(1 + \tau_c) \cdot y$ (pre-tax) to $(1 + \tau_c) \cdot c$ (post-tax) by subtracting actual taxes paid and actual transfers received.

[52] If $y - t_y$ and g are both cash, then they contribute to t_c in proportion. If g is a pure in-kind transfer such as health insurance that faces no consumption tax, then t_c would be assigned fully to $y - t_y$.

[53] For linear taxes $t_y = \tau \cdot y$ and $t_c = \tau_c \cdot c$, we have $(1+\tau_c)c = y \cdot (1-\tau) + g$ so that $c = y \cdot (1-\tau)/(1+\tau_c) + g/(1+\tau_c)$. Hence $y - t_y - t_c = y \cdot (1-\tau)/(1+\tau_c)$ and $g_n = g/(1+\tau_c)$. Hence τ and τ_c add up to the standard $(\tau+\tau_c)/(1+\tau_c)$.

[54] We can still have high tax rates for individuals with no income and no transfers who consume through dissaving, but this issue is typically alleviated when aggregating by income groups.

[55] As an illustration, suppose half of workers produce a private good and half produce an untaxed public good funded by the government. All workers are identical and paid the same and hence the tax take in this economy is intuitively 50%. Suppose the government uses only a consumption tax on the private good (i.e $T_y = 0$ and $T_c = T_{cy} = Y/2$). The consumption tax rate has to be 100% to fund the public good production, i.e. the after-tax price of the private good is twice the pre-tax production price of the private good. In this case our proposed measure $T_{cy}/Y = 50\%$ gets it right but the traditional measure is $T_c/(Y + T_c) = .5/1.5 = 1/3$. In basic-price national income, the public sector is half of the economy but in traditional national income, the public sector is only 1/3 of the economy because the private sector gets a heavier weight due to the consumption tax it faces.

Taxes paid by nonprofits. Some nonprofit organizations pay capital taxes: corporate taxes on the profits of the companies they invest in, property taxes on the assets the own. To the extent that nonprofits provide collective wealth and services, they should be left out of distributional analysis. To match national income, both their primary capital income and the corresponding taxes should be allocated in a distributionally-neutral manner, i.e., proportionally to after-tax disposable income.

A.3 Illustration of the Current-Tax Method: Case Studies

We illustrate our current-tax methodology with case studies in the year 2018. These case studies only use publicly available information, and are summarized in Table A2.

Jeff Bezos. Start with Jeff Bezos, the richest person in the United States in 2018 according to *Forbes*. To compute his tax rate, we need to estimate his pre-tax income from all sources and the taxes he paid (directly and indirectly) worldwide.

Bezos derives most of his income from his stake in Amazon. As reported in its annual 10-K report to the Securities and Exchange Commission (SEC), the company made $11.3 billion in pre-tax income globally in 2018.[56] Since Bezos owned 16.3% of Amazon, he earned 16.3% of Amazon's profit, i.e., around $1.84 billion. Even though Amazon did not pay dividends in 2018, its profits did constitute income for Amazon's shareholders like Bezos—income that was fully saved and reinvested in the firm. Bezos also earned income from other investments, such as his stake in the *Washington Post*. Public sources suggest he earned around $250 million in taxable income from these other investments.[57] We disregard other income sources such as imputed rents on real estate properties and income earned on pension assets and trusts, which are second-order for our purposes.

Bezos also realized capital gains by selling Amazon stocks, $33 million according to SEC form 4 public reports. Since he founded Amazon, and prior to 2018 Amazon made little profit, Bezos's cost basis was small in 2018. Virtually all of his realized capital gains reflected pure asset price appreciation, not the effect of past or current retained earnings (already included in income). Therefore we include the $33 million in realized capital gains in Bezos's income. Because his realized capital gains are small, including these gains in income makes negligible difference to Bezos's effective tax rate.

[56] This number is net of interest and depreciation; it is conceptually close to corporate profits as included in national income. There are differences between profit accounting in financial statements and in the national accounts; for our purposes in this paper, however, these differences are second-order.

[57] According to ProPublica (Eisinger et al., 2021), Bezos reported $284 million in total income on his individual income tax return. Of this, $1.7 million corresponds to Amazon compensation ($81,840 in wage and $1,600,000 in other compensation—security detail—according to public SEC forms). Since Amazon did not distribute dividends and since according to SEC form 4 public reports Bezos realized $33 million in capital gains by selling Amazon stocks (see below), around $250 million in income derived from non-Amazon holdings.

We compute Bezos's total income tax as his share of the income taxes paid by Amazon plus the income taxes he paid directly. Amazon paid $1.18 billion in cash income taxes in 2018 to federal, state, and foreign governments combined, an effective tax rate of 10.5%.[58] In our methodology, Bezos paid $193 million in corporate taxes, namely his share (16.3%) of Amazon's corporate income taxes (or, equivalently, 10.5% of his Amazon income of $1.84 billion). Moreover, Amazon paid business property taxes. The amounts are not publicly disclosed. We can estimate these taxes as roughly equal to 1% of Amazon's capital stock, the US-wide average business property tax rate. This adds around $100 million in taxes for Bezos. Last, according to ProPublica, Bezos paid $43 million in federal individual income taxes. As a resident of Washington State, Bezos did not pay state income taxes. Other taxes paid by him directly or through Amazon are negligible for our purposes.[59] His total tax payments thus amounted to $337 million, an effective tax rate of 15.2%.[60]

Two remarks about this result are worth mentioning. First, if we focus on federal taxes alone (as US government agencies do), Bezos's effective tax rate was only 1.9%. According to its 10-K, Amazon did not pay any federal corporate income taxes in 2018. Property taxes are all paid to state, local, and foreign governments. The only federal tax Bezos paid was the individual income tax. For many economic questions (e.g., the study of behavioral response to taxes) the relevant tax rates are those including all levels of governments. However in some contexts (e.g., policy discussions of federal tax reforms), effective federal tax rates can also be relevant. Second, our methodology to allocate the corporate tax implies a higher effective tax rate for Bezos than the conventional approach. In the conventional approach the amount of corporate tax allocated to Bezos has nothing to do with the amount paid by Amazon, since the corporate tax is shifted to workers and capital owners nationally. In the CBO methodology, Bezos pays about $41 million in corporate taxes, as opposed to $193 million with our methodology.[61]

Warren Buffett. Buffett's situation is to some extent similar to Bezos. The company he owns—Berkshire Hathaway—does not distribute dividends and he realizes little capital gains. Thus the bulk of his taxes correspond to his share of Berkshire Hathaway's corporate and

[58] Using provisions for income taxes instead of cash income taxes paid gives a similar effective tax rate, 10.6%. Both measures have merits and demerits. One issue with provisions for income taxes paid is that these provisions include tax contingencies—taxes that have not been paid but that companies estimate have a more than 50% chance to be eventually paid as a result of audits and other enforcement activities. Because some of these tax contingencies end up not being paid (e.g., due to a lapse in statute of limitation), provisions for income taxes can over-estimate actual tax payments.

[59] Residential property taxes paid by Bezos are likely to be negligible compared to his income. Sales taxes paid by Bezos are likely to be negligible too. For example, in the case where he consumed $10 million of taxable goods in Seattle, the associated sales tax would be $1.025 million (6.5% rate in Washington state plus 3.75% rate in Seattle), increasing his effective tax rate by only 0.05 percentage point.

[60] This effective rate is equal to $337 million divided by $2.2 billion in income: the $100 million in Amazon property taxes have to be added to the income denominator because they are not counted as income in corporate income statements.

[61] Specifically, in 2018 total U.S. corporate tax revenues (federal plus state) added up to $283 billion. Bezos

property taxes. According to the conventional approach Buffet pays essentially zero tax, but in our methodology his effective tax rate (taking into account all taxes paid) was 18.4% in 2018.

Specifically, Buffett had $8.2 billion in income, corresponding to his share (30.2%) of Berkshire Hathaway's $27.0 billion in pre-tax profit.[62] According to ProPublica, and consistent with public SEC reports, Buffett had negligible reported individual income ($24.8 million, out of which he paid $5.36 million in federal taxes). According to its 10-K, Berkshire Hathaway's effective corporate global cash income tax rate was 16.1%—and 18.4% when adding our estimate of business property taxes. Since Buffett had negligible individual taxable income, the taxes he paid at the individual level were negligible relative to his share of the taxes paid by Berkshire Hathaway. Buffett's effective tax rate was thus equal to Berkshire Hathaway's, 18.4%.

Buffett's case illustrates that the corporate tax—and to a lesser extent business property taxes—serve as a backstop for the ultra-wealthy. Without these taxes, Buffett's effective tax rate would be 0% out of $8.2 billion in income. Moreover, like for Bezos, our methodology assigns much more corporate tax to Buffett than the conventional approach. Since Buffett has negligible individual taxable income, in the CBO methodology Buffett is assigned virtually no corporate tax, even though Berkshire Hathaway, of which he owns 30%, paid more than $4 billion in cash corporate income taxes in 2018. A complete quantification of the taxes paid by the top 400 wealthiest Americans, systematically linking businesses to owners using administrative data, is presented in Balkir et al. 2025.

A.4 Distributional Analysis in a Two-Sector Capital and Labor Model

In this appendix section, we contrast our analysis with the traditional incidence analysis in the two-sector model from Section 4.2.

Equilibrium of the model. The following three equations determine the equilibrium (w, r, k):

$$r = f'(k), \quad w = f(k) - kf'(k), \quad k = k(r \cdot (1 - \tau_K)). \tag{8}$$

This simple model has the advantage of being representable in a standard capital demand and supply diagram (Figure A3). Even though this is a general equilibrium model, the diagram is the same as the standard textbook one-market model of tax incidence. The demand for capital is $r = f'(k)$ and is downward sloping (as $f''(k) < 0$). The supply for capital is $k =$

earned 0.00002% of all reported taxable wages and 0.02% of all reported taxable capital income (dividends, taxable interest, net rents and royalties, and realized capital gains); hence he gets assigned 25% × 0.00002 + 75% × 0.02 = 0.0145% of corporate tax payments (or $41 million) if one applies the CBO methodology.

[62] Berkshire Hathaway's pre-tax profits are computed as pre-tax income as officially reported in the 10-K ($4.0 billion), plus realized gains on investments ($22.5 billion), plus (imputed) business property taxes paid ($0.6 billion, computed like in the case of Amazon as 1% of net property and equipment). Unrealized gains on investments are removed because they are not taxable and not part of conventionally defined income. Consistently, we measure income tax paid as cash tax paid (as in the case of Amazon).

$k(r \cdot (1-\tau_K))$ and is upward sloping (and flat when $e_K = \infty$). The surplus accruing to workers is $w = f(k) - rk = \int_0^k f'(\kappa)d\kappa - rk$ and can be read off as the area below the demand curve and above the horizontal line at r. The surplus accruing to capitalists is the area above the supply curve and below the horizontal line at $\bar{r} = r \cdot (1-\tau_K)$. Capital taxes are the rectangle $(r-\bar{r})k$. The triangle pointing toward the no-tax equilibrium $f'(k^*) = r^*, k(r^*) = k^*$ is the usual deadweight burden. It is equal to the loss in surplus of workers and capitalists created by the tax τ_K over and above its revenue yield $(r-\bar{r})k$.

Distributional current-tax analysis. How should we describe such an economy in a current-tax distributional table? wL is labor income and rK is capital income, as would be measured in national accounts statistics. While it is true that τ_K affects w negatively, w is the actual pre-tax wage rate in the economy. Similarly, τ_K affects r positively, but the actual pre-tax rate of return is r and not the lower r^*. The logical description of current pre-tax income, post-tax incomes, and tax paid is thus the following. On the labor side, pre-tax labor income is wL, post-tax labor income is $\bar{w} = w(1-\tau_L)L$, and workers pay $\tau_L wL$ in taxes. On the capital side, pre-tax capital income is rK, post-tax capital income is $\bar{r}K = r(1-\tau_K)K$, and capitalists pay $\tau_K rK$ in taxes.

Contrast this with the distributional tax analysis carried out by US government agencies. This analysis ignores the deadweight burden and considers that capital taxes $\tau_K rK = (r-\bar{r})K$ are shared by capitalists who pay $(r^*-\bar{r})K$, and by workers who pay $(r-r^*)K$. The pre-tax income of workers is $wL + (r-r^*)K$ and the pre-tax income of capitalists is $rK + (r^*-\bar{r})K = (r^* - \tau_K r)K$. These concepts are neither the actual incomes going to workers and capitalists before tax, nor the incomes that would go to workers and capitalists absent taxes (since the change in K and deadweight burden are ignored). This might be a defensible assumption for small taxes, where deadweight burden is second order. In practice, however, taxes are large. If the supply of capital is perfectly elastic, then the capital tax is borne fully by labor. In conventional distributional analysis, it is equivalent to a tax on inelastic labor, even though the two taxes have drastically different efficiency implications.

Tax incidence analysis. We consider a small increase in the capital tax rate $d\tau_K$ and trace out its effects dk, dr, dw. Differentiating the 3 equations in (8), we have two equations on the production side:

$$\frac{dk}{k} = \sigma \cdot \left[\frac{dw}{w} - \frac{dr}{r}\right], \quad dw + k \cdot dr = 0.$$

The first equation is the definition of the elasticity of substitution between labor and capital, σ.[63]

The second equation is obtained by differentiating $f(k) = rk + w$ and using $f'(k) = r$. This

[63]With Cobb-Douglas production functions of the form $F(K,L) = A \cdot K^\alpha L^{1-\alpha}$ then $\alpha = rK/Y$ is constant

equation is key: it states that the effects of the reform on factor prices sum to zero. What labor loses due to reduced wages is exactly what capital gains through a higher return.

On the supply side, we have:

$$\frac{dk}{k} = e_K \cdot \frac{d\bar{r}}{\bar{r}} = e_K \cdot \left(\frac{dr}{r} - \frac{d\tau_K}{1-\tau_K} \right).$$

Combining and rearranging, and denoting by $\alpha = rK/Y = rk/f(k)$ the share of capital income in the economy and hence $1-\alpha = w/f(k)$ the labor share, we obtain:

$$\frac{dr}{r} = \frac{(1-\alpha)e_K}{(1-\alpha)e_K + \sigma} \cdot \frac{d\tau_K}{1-\tau_K}, \quad \frac{dk}{k} = -e_K \cdot \frac{\sigma}{(1-\alpha)e_K + \sigma} \cdot \frac{d\tau_K}{1-\tau_K}, \quad dw = -kdr.$$

These equations display the usual lessons from tax incidence. First, if $\sigma = \infty$, then a capital tax increase has no effect on factor prices r and w. It only affects capital through a pure supply-side response: $dk/k = -e_K d\tau_K/(1-\tau_K)$. Second, if $\sigma < \infty$, then capital supply responses affect factor prices, spreading partly the incidence of the tax onto wages. The shift to wages is small whenever e_K is small relative to σ.

This is illustrated in Figure A4. The increase in τ_K shifts the equilibrium. The reduction in $\bar{r}$ along the supply curve is attenuated by an increase in dr along the demand curve. The response dk is attenuated relative to the case where r is fixed. Capital tax revenue is $\tau_K rk = (r-\bar{r})k$. Its change can be decomposed into three terms depicted on the graph:

$$d[(r-\bar{r})k] = -kd\bar{r} + kdr + (r-\bar{r})dk. \tag{9}$$

The first term $-kd\bar{r} > 0$ is the direct effect due to a lower net-of-tax rate of return $\bar{r}$. The second term $kdr > 0$ is due to a higher pre-tax rate of return r. Importantly, this term is exactly equal to $-dw$, i.e., what is lost by workers due the reduction in the wage rate w. The third term is the tax revenue lost due to the supply-side response of capital (itself triggered by $d\bar{r}$). This tax revenue loss is equal to the increase in the deadweight burden triangle of the tax.

Optimal tax analysis. Suppose the social marginal welfare weight on capitalists is zero. Maybe capitalists are much more well-off than workers (and hence have much lower marginal utility), or maybe all residents are workers and the country attracts capital from abroad only. In this case, society sets τ_K to maximize workers' income $w + (r-\bar{r})k$ where w is the wage and $(r-\bar{r})k$ is the tax collected from capitalists. As $w + rk = f(k)$, social welfare is $w + (r-\bar{r})k = f(k(\bar{r})) - \bar{r}k(\bar{r})$. The government effectively chooses $\bar{r}$ along the supply side curve $k(\bar{r})$ to maximize surplus—the area above the line $\bar{r}$ and below the demand curve for capital (blue area in Figure A3). The first-order condition for the optimum τ_K is such that:

$$0 = (f'(k) - \bar{r})dk - kd\bar{r} = -kd\bar{r}\left[1 - \frac{r-\bar{r}}{\bar{r}}\frac{\bar{r}}{k}\frac{dk}{d\bar{r}}\right] = -kd\bar{r}\left[1 - \frac{\tau_K}{1-\tau_K}e_K\right]. \tag{10}$$

This leads to the usual inverse-elasticity rule optimal tax rate $\tau_K^* = 1/(1+e_K)$.

The key insight is that the optimal tax rate only depends on the supply elasticity e_K, not on whether the tax on capital is shifted to workers. In other words, the supply elasticity is a sufficient statistics for the optimal tax rate (and the elasticity of substitution σ is irrelevant). The intuition for this result can be seen on Figure A4. Workers' welfare is the wage area w plus the tax rectangle. When τ_K increases, the reduction in wages dw is fully offset by the increase in tax revenue kdr. As a result, the tradeoff is only about the mechanical increase in tax revenue $kd\bar{r}$ vs. the revenue loss due to the supply side response $(r-\bar{r})dk$ (depicted in blue shaded areas in Figure A4).[64] Intuitively, setting τ_K is equivalent to setting $\bar{r}$ so that the implicit changes in r triggered by τ_K can be neutralized.[65] This result is a special case of a more general result first derived by Diamond and Mirrlees (1971).[66] Optimal tax formulas can be expressed solely in terms of social marginal welfare weights and household level responses: supply elasticities (for production factors such as capital or labor) or demand elasticities (for consumption goods). Conditional on these, elasticities of substitution within the production sector (e.g., between capital and labor) are irrelevant.

Distributional tax-reform analysis. Because the effects of taxes on prices do not matter normatively, we recommend ignoring them for distributional tax-reform analysis. Consider a capital tax increase. On the equity side of the trade-off, the relevant impact is the direct effect of the tax on capital owners, ignoring both supply-side responses and any effects on pre-tax wages and rates of return. For example in the case of a corporate tax increase, all that matters is the mechanical changes in corporate tax payments by income group, which can be computed using the current-tax table (reporting how much corporate tax each group of the population pays today). The welfare costs of these direct effects can be aggregated across income groups using social marginal welfare weights. On the efficiency side of the trade-off, the sufficient statistic is the total change in tax revenue due to supply-side responses, ignoring again any price effects. The revenue change does not need to be distributed by groups. We provide a concrete illustration in the case of an increase in the US corporate income tax rate in Section 5.1 below.

---

and $\sigma = 1$. With a CES production function $F(K,L) = [\mu K^{(\sigma-1)/\sigma} + (1-\mu)L^{(\sigma-1)/\sigma}]^{\sigma/(\sigma-1)}$ the elasticity of substitution σ is constant.

[64] The derivation has been made (independently) by Piketty (2000) and Mankiw (2001) in the special case where $e_K = \infty$ (horizontal supply curve in our diagrams) as a way to demonstrate the uselessness of capital taxes in the standard model in which the infinite capital supply elasticity arises from infinite horizon utility maximization and $\bar{r}$ is pinned down by the exogenous discount rate δ. This derivation based on long-run outcomes is distinct from the classical Chamley-Judd zero capital tax result (see Saez and Stantcheva, 2018; and Straub and Werning, 2020).

[65] This result carries over more generally even if government puts a weight on capitalists (say $g_K < 1$ per \$ of capitalist surplus lost). The reform depicted on Figure A4 reduces the surplus of capitalists by $kd\bar{r} < 0$ so that the optimum first-order condition simply becomes $(r-\bar{r})dk = (1-g_K)kd\bar{r}$ (instead of $(r-\bar{r})dk = kd\bar{r}$) leading to the classic optimal tax formula $\tau_K^* = (1-g_K)/(1-g_K+e_K)$.

[66] Piketty and Saez (2013) and Saez and Stantcheva (2018) show how it applies to inheritance taxation and capital income taxation respectively.

A capital tax increase also affects factor prices and the distribution of pre-tax income. It reduces workers' wages and increases capitalists' pre-tax income, typically leading to an increase in overall income inequality. It also changes the amount of taxes paid by each group. But these effects are normatively irrelevant because all the pre-tax price effects can be neutralized by a corresponding adjustment of all the other taxes which is budget neutral. Of course, this result arises in the context of the specific neo-classical model of Diamond and Mirrlees (1971). It is important to note, however, that conventional distributional tax analysis considers the very same type of models.[67]

Contrasting Tax Incidence vs. Optimal Tax in the Two-Sector Model. Let us formally contrast the tax incidence approach with the optimal tax approach in a slightly extended version of the two-sector model labor and capital model to allow for elastic labor supply and the presence of inactive benefit recipients.

A population of size 1 is divided between p_L workers, p_K capitalists, and $p_0 = 1 - p_L - p_K$ inactive benefit recipients. The government raises revenue with taxes on domestic labor income and capital income at flat rates τ_L and τ_K and uses it to fund a lumpsum transfer R to all. Workers have all identical individual utilities of the form $u^L(c, l) = c - l^{1+1/e_L}/(1+1/e_L)$ (where c is consumption and l is labor supply) which they maximize under the budget constraint $c = \bar{w} \cdot l + R$ where $\bar{w} = w(1 - \tau_L)$ is the net-of-tax wage rate. The first order condition $l^{1/e_L} = \bar{w}$ generates an isoelastic labor supply $l = \bar{w}^{e_L}$ which aggregates into macro-level labor supply $L = p_L \cdot l = L(\bar{w})$ with elasticity e_L. Recall that capitalists choose to invest a part k of their total capital k_0 at home with rate of return $\bar{r} = r(1 - \tau_K)$ and the remaining part $k_0 - k$ abroad with a rate of return r_0. They have a money metric utility with home-bias $u^K(c, k) = c + a(k)$ with $c = \bar{r}k + R + r_0(k_0 - k)$, leading to a first order condition $a'(k) = r_0 - \bar{r}$ which defines an upward sloping supply of aggregate domestic capital $K = p_K k = K(\bar{r})$ with elasticity e_K. The inactive have utility $u^0(c) = c$ and simply consume the lumpsum grant with $c = R$.

The following four equations determine the equilibrium (w, r, K, L) of the model as a function of the tax rates τ_L, τ_K, the production function $F(.,.)$ and the supply functions $L(.), K(.)$:

$$r = F_K(K, L), \quad w = F_L(K, L), \quad L = L(w \cdot (1 - \tau_L)), \quad K = K(r \cdot (1 - \tau_K)). \tag{11}$$

Optimal tax. Let us start with the optimal tax approach. The government chooses τ_L, τ_K, R to maximize social welfare

$$SW = p_L g_L u^L + p_K g_K u^K + p_0 g_0 u^0,$$

with g_L, g_K, g_0 the exogenous social marginal welfare weights on each group which we assume

[67] In the real world, taxes can have effects that are more complex than what is captured by neoclassical models, in which case price effects may not be irrelevant. We extend the analysis along those lines in Section A.5.

average to one (without loss of generality) so that

$$SW = R + p_L g_L \cdot [\bar{w}l - l^{1+1/e_L}/(1+1/e_L)] + p_K g_K \cdot [\bar{r}k + r_0(k_0 - k) + a(k)].$$

The government budget constraint is:

$$R = \tau_L wL + \tau_K rK = (w - \bar{w})L + (r - \bar{r})K = F(K, L) - \bar{w}L - \bar{r}K$$

which can be plugged in the social welfare function. Hence, equivalently, the government choose $\bar{w}$ and $\bar{r}$ to maximize:

$$SW = F(K(\bar{r}), L(\bar{w})) - \bar{w}L(\bar{w}) - \bar{r}K(\bar{r}) + p_L g_L[\bar{w}l - l^{1+1/e_L}/(1+1/e_L)] + p_K g_K[\bar{r}k + r_0(k_0 - k) + a(k)].$$

Importantly, pretax prices w and r have disappeared from the objective function. The government can use taxes τ_L and τ_K to determine the after-tax prices $\bar{w}$ and $\bar{r}$ ignoring the effects on pre-tax prices, one of the key results from Diamond and Mirrlees (1971). Using the envelope conditions that l and k choices maximize individual utilities, and using that $F_K = r$ and $F_L = w$, we obtain the following first order condition for government optimization:

$$0 = \frac{dSW}{d\bar{r}} = (r - \bar{r})\frac{dK}{d\bar{r}} - K + p_K g_K k = \frac{r - \bar{r}}{\bar{r}} e_K K - K + g_K K.$$

$$0 = \frac{dSW}{d\bar{w}} = (w - \bar{w})\frac{dL}{d\bar{w}} - L + p_L g_L l = \frac{w - \bar{w}}{\bar{w}} e_L L - L + g_L L.$$

These two equations lead to the standard optimal tax formulas:

$$\frac{\tau_K^*}{1 - \tau_K^*} = \frac{r - \bar{r}}{\bar{r}} = \frac{1 - g_K}{e_K} \quad \text{i.e.} \quad \tau_K^* = \frac{1 - g_K}{1 - g_K + e_K},$$

$$\frac{\tau_L^*}{1 - \tau_L^*} = \frac{w - \bar{w}}{\bar{w}} = \frac{1 - g_L}{e_L} \quad \text{i.e.} \quad \tau_L^* = \frac{1 - g_L}{1 - g_L + e_L}.$$

Optimal tax rates depend solely on the supply side behavioral responses of labor and capital e_L and e_K along with the social welfare weights that the government assigns to each group g_L and g_K. Tax incidence on pretax prices is irrelevant because it affects the splitting of production into pretax labor and capital income: $F(K, L) = wL + rK$ but what matters for the government budget is total resource $F(K, L)$ and what matters for individuals are aftertax prices.

Tax incidence. Let us now consider the tax incidence approach starting from a given tax system (τ_L, τ_K). We consider a small increase in the capital tax rate $d\tau_K > 0$ and trace out its effects dK, dL, dr, dw. Differentiating the 4 equations in (11), we have two equations on the production side:

$$\frac{dK}{K} - \frac{dL}{L} = \sigma \cdot \left[\frac{dw}{w} - \frac{dr}{r}\right], \quad L \cdot dw + K \cdot dr = 0.$$

The first equation is the definition of the elasticity of substitution between labor and capital, σ. The second equation is obtained by differentiating $F(K,L) = rK + wL$ and using $F_K = r$ and $F_L = w$.

On the supply side, we have two equations:

$$\frac{dK}{K} = e_K \cdot \frac{d\bar{r}}{\bar{r}} = e_K \cdot \left(\frac{dr}{r} - \frac{d\tau_K}{1-\tau_K} \right), \quad \frac{dL}{L} = e_L \cdot \frac{d\bar{w}}{\bar{w}} = e_L \cdot \frac{dw}{w}.$$

Combining and rearranging, on the capital side we obtain:

$$\frac{dr}{r} = \frac{(1-\alpha)e_K}{(1-\alpha)e_K + \alpha e_L + \sigma} \cdot \frac{d\tau_K}{1-\tau_K}, \quad \frac{d\bar{r}}{\bar{r}} = -\frac{\alpha e_L + \sigma}{(1-\alpha)e_K + \alpha e_L + \sigma} \cdot \frac{d\tau_K}{1-\tau_K},$$

and on the labor side:

$$\frac{d\bar{w}}{\bar{w}} = \frac{dw}{w} = \frac{-\alpha e_K}{(1-\alpha)e_K + \alpha e_L + \sigma} \cdot \frac{d\tau_K}{1-\tau_K}.$$

Therefore, pretax price incidence shifts the initial capital tax increase partly onto labor: the after-tax return on capital falls by less than the new tax but the after-tax wage also falls. Hence, in the optimal tax approach discussed just above where the government optimizes $\bar{r}$ and $\bar{w}$, $d\tau_K > 0$ amounts to reducing $d\bar{r}$ by less than $-rd\tau_K$ but at the same time reducing $\bar{w}$ by $d\bar{w}$. Therefore, it mixes a (smaller) tax increase on capital with a tax increase on labor. The welfare effects of the reform $d\tau_K$ amount to analyzing the welfare effects of $d\bar{r}$ and $d\bar{w}$ and ignoring the irrelevant price effects as discussed above.

If the labor tax is optimal and equal to τ_L^*, then $d\bar{w}$ has zero first order welfare effects, and hence the welfare effects of $d\tau_K > 0$ are the same as the welfare effects of $d\bar{r} < 0$. If $\tau_K < \tau_K^*$, increasing the tax rate is desirable whether or not price effects are taken into accounts.

If the labor tax is suboptimal $\tau_L < \tau_L^*$ then $d\bar{w} < 0$ has a positive first order welfare effect. Therefore, if $\tau_K < \tau_K^*$, then $d\tau_K > 0$ is desirable both because it increases the tax on capital and also because it implicitly increases the tax on labor.

However, if the labor tax is too large $\tau_L > \tau_L^*$ then $d\bar{w} < 0$ has a negative first order welfare effect. Therefore, if $\tau_K < \tau_K^*$, then $d\tau_K > 0$ will be desirable if and only if the positive impact of $d\bar{r} < 0$ is larger than the negative impact of $d\bar{w} < 0$. Which effect dominates depends on which tax rate is furthest away from its optimum. If τ_K is only slightly below τ_K^* and τ_L is substantially above τ_L^*, then the $d\bar{w}$ welfare effect will dominate making the reform $d\tau_K > 0$ undesirable.

While it is certainly important for a policy maker to learn from classic tax incidence that a reform $d\tau_K > 0$ may be undesirable even if $\tau_K < \tau_K^*$, it is also important for economic advice to explain that the reason $d\tau_K > 0$ is not desirable is because τ_L is too low and that combining an even greater capital tax increase with a reduction of τ_L can achieve the goal of policy maker. This is why we view classic tax incidence as useful but overly narrow and why we think that

optimal tax analysis offers a vital broader picture view for the analysis of tax reform. Put simply, the optimal tax approach tells the policy maker which direction to go; the tax incidence analysis can provide the technical pathway on how to get there.

Generalization: Consumption taxes. Let us consider the basic supply and demand tax incidence diagram for one good from introductory economics, as illustrated on appendix Figure A5(a). Formally, the producer profit is $\Pi = pQ - c(Q)$ where p is the pre-tax price of the good, Q the quantity produced, and $c(Q)$ the increasing and convex cost of producing a quantity Q. Profit maximization implies that $p = c'(Q)$ which defines the supply curve $S(p)$. The consumer utility is $V = v(Q) - \bar{p}Q$ where Q is the quantity of the good consumed, $v(Q)$ the increasing and concave utility of consuming Q, and $\bar{p} = p + t$ the after-tax price of the good (with t the tax per unit of good). Utility maximization implies $v'(Q) = \bar{p}$ which defines the demand curve $D(\bar{p})$. The key point is that, in the Diamond and Mirrlees model, pure profits are assumed to be fully taxed away.[68] Therefore, taxes collected are $T = tQ + \Pi = \bar{p}Q - c(Q)$.

The classic Ramsey tax problem sets tax rates to collect a certain tax revenue while minimizing utility loss. Therefore, the key tradeoff is consumer surplus vs. taxes collected at the margin. As illustrated on appendix Figure A5(b), increasing the tax t mechanically increases tax revenue (and correspondingly reduces consumer surplus) but it also reduces taxes through the behavioral response (and correspondingly increases deadweight burden). Because pure profits are in the tax base, the increase in tax from the consumption good due to $dp < 0$ is fully offset by the loss of profit $d\Pi$ and hence this margin is irrelevant.[69]

Mathematically, the Lagrangian takes the form

$$V + \lambda T = v(D(\bar{p})) - \bar{p}D(\bar{p}) + \lambda[\bar{p}D(\bar{p}) - c(D(\bar{p}))].$$

Hence, (and using the envelope conditions $v'(Q) = \bar{p}$, $c'(Q) = p$), the first order condition in $\bar{p}$ takes the form: $-D(\bar{p}) + \lambda D(\bar{p}) + \lambda D'(\bar{p})[\bar{p} - p] = 0$ which can be rewritten as the classic inverse elasticity formula:

$$\frac{t}{p+t} = \frac{1}{\varepsilon_D} \cdot \frac{\lambda - 1}{\lambda}, \tag{12}$$

with $\varepsilon_D = -\bar{p}D'(\bar{p})/D(\bar{p}) > 0$ the elasticity of demand for the good from the consumer and $\lambda > 1$ reflecting the fact that the marginal dollar of tax creates a welfare loss in excess of one dollar on the consumer. The elasticity of supply coming out of the production side does not appear in equation (12).

With only one taxed good, the Ramsey problem is not meaningful but it is straightforward to consider multiple goods. With separability $V = \sum_i v_i(Q_i) - \bar{p}_i Q_i$, the demand function for

[68] Pure profits arise in this simple one good model but would not exist in a model with several production factors and constant returns to scale (as in the labor and capital model discussed above).

[69] This is of course the same logic as in the two-factor model where the lost wages dw were made up by more

each good Q_i depends only on its own price $\bar{p}_i$, and the same analysis carries through and equation (12) applies to each good with the same λ, which is the basic Ramsey inverse elasticity rule.[70]

A.5 Incorporating Non-Standard Behavioral Effects

In the neoclassical models considered so far, distributional tax-reform analysis is straightforward to conduct. These models also have the advantage of accommodating tax avoidance responses which are often first order (Slemrod, 1995). Their main limitation is that they do not allow for some important non-standard behavioral responses to taxes uncovered by the modern empirical literature on tax incidence. In this Section we take stock of this body of work. We show that non-standard incidence can be incorporated into our tax-reform framework, and provide an application to a reform that would replace employer-provided health insurance premiums by a payroll tax.

A.5.1 Non-Standard Incidence: Lessons from the Recent Literature

We define as non-standard incidence any incidence effect that cannot be reconciled with the neoclassical model used above. Table A3 provides a summary of the key non-standard behavioral responses to taxes uncovered by the recent literature, tax by tax.[71] Benzarti (2025) provides a comprehensive recent survey of tax incidence anomalies.

Corporate taxes. A number of papers find non-standard effects of the corporate tax that operate through bargaining over the distribution of value-added within businesses. Kennedy et al. (2022) show that the large 2018 cut in the US corporate tax rate also generated earnings gains for workers in treated C-corporations relative to workers in control S-corporations. This contradicts classical incidence where wages should adjust across the board. Furthermore, the wage gains are concentrated among top 10% and especially top 1% earners with no gain for the bottom 90%. Highly paid workers capture 32% of the corporate tax cut (Table 11 in Kennedy et al., 2022).[72] In Table A3, we use Kennedy et al.'s estimates and assign 2/3 of a corporate

capital income kdr.

[70] Ramsey (1927) did not assume that pure profits could be taxed so that Ramsey's formulas do depend on supply elasticities as well. However, Diamond and Mirrlees (1971) noted that constant returns to scale, which rules out pure profits, is a better assumption in general equilibrium. Hence, the standard assumption in modern optimal tax theory has been to assume that there are no pure profits or that they can be taxed away fully. Stiglitz and Dasgupta (1971) is the classic reference exploring this point.

[71] The individual income tax does not exhibit major non-standard incidence effects, except for the fact that individuals do not have perfect understanding of the tax system (see e.g. Rees-Jones and Taubinsky (2020)).

[72] Fuest, Peichl, and Siegloch (2018) show that municipality-level corporate tax cuts in Germany also affect wages, with workers receiving about 40% of the tax windfall. This suggests that bargaining power within the firm affects how a corporate tax windfall is distributed, with strong unions in Germany perhaps able to spread windfalls more equitably among workers.

tax change to profits and 1/3 to workers. Crucially, these within-firm effects have nothing to do with the macroeconomic effect of taxes on factor prices in classical incidence models, and hence are relevant for assessing the direct welfare effects of a reform.[73]

Consumption taxes. The standard model predicts that increases vs. decreases in taxes should have symmetric effects. This result is strikingly proven false by Benzarti et al. (2020) in the case of the value-added tax (VAT), the major form of consumption tax worldwide, using a comprehensive analysis of VAT reforms in Europe from 1996 to 2015. While producers can pass almost all of a VAT increase onto consumers, VAT cuts are only half passed onto consumers and hence benefit businesses–and their workers and suppliers. These asymmetric price effects persist several years after VAT changes take place. The most likely explanation is that businesses can justify a price increase if there is a tax increase, but can silently pocket a tax decrease with inattentive consumers.[74] This asymmetric evidence is based on many VAT changes in Europe and hence solidly established. For distributional tax reform analysis, this implies that a VAT tax increase can be assigned to the corresponding consumers as in conventional analysis, but a VAT tax decrease should be shared half between consumers and half for businesses and their workers.

There is more uncertainty on how the tax windfall going to businesses should be split between profits and workers. The estimates from Benzarti and Carloni (2019) for a single specific VAT cut for restaurants in France show that of the incidence on businesses, 75% goes to profits and 25% to workers but it is hard to know whether such numbers generalize to other sectors or countries as they likely depend on workers' bargaining institutions and power.[75] Hence, in our summary Table A3, we split a VAT cut 50% to consumers, 37.5% to profits, and 12.5% to workers with the latter two figures being highly uncertain. Because some of the VAT reforms considered by Benzarti et al. (2020) are very sector specific (e.g., hairdressers or restaurants), we conjecture in Table A3 that similar effects would hold for excise taxes as such taxes are like the VAT typically built in the price posted to consumers.

The persistent asymmetric result by Benzarti et al. (2020) also shows that there is no single equilibrium, since a cut followed by an offsetting increase in VAT rate on a specific good seem to lead to permanently higher prices for the good (Figure 2 in Benzarti et al. 2020 provides a striking case study for hairdressers vs. beauty salons in Finland). This radical departure from

[73]We did not incorporate these effects in our corporate tax reform analysis of Section 5.1, because the wage effect in the United States is highly concentrated at the top of the distribution, so that accounting for it has only minor effects relative to assuming that the full impact is on profits (see Kennedy et al. 2022, for a detailed analysis).

[74]See Kosonen (2015) and Harju et al. (2018) for a more detailed discussion in the context of the hairdressing and restaurant industries in Finland showing that non-standard incidence is concentrated among smaller businesses.

[75]Benzarti and Carloni (2019) find that firm owners pocketed around 55.7 percent of the VAT cut and employees received 18.6 of the VAT cut, making for a 3/4 vs. 1/4 split between profits and workers in this case.

equilibrium analysis means that the no-tax counterfactual of classical tax incidence analysis might not even be well defined, further supporting our current-tax approach that does not rely on such a counterfactual.

US sales taxes are not visible on posted prices and charged at the checkout. Empirical work shows that they are passed to consumers symmetrically (for cuts or increases) and generally fully (see e.g. Poterba 1996 and Besley and Rosen 1999 for empirical studies). Chetty, Looney and Kroft (2009) show that consumers also under-react to changes in sales taxes relative to changes in excise taxes that are included in posted prices, a relevant finding to inform distributional tax-reform tables that we point out in Table A3.

Payroll taxes. A celebrated result in classical incidence analysis is that employer and employee payroll taxes are equivalent. In the real world, this result fails to materialize. A number of studies compellingly show that *employee* payroll taxes changes affect the net wage earnings of the corresponding workers one-to-one but that employers fail to pass changes in *employer* payroll taxes to the corresponding workers, likely because of wage rigidities.[76] As a result, an increase in employer payroll taxes likely reduces wages across the board and probably profits inside the business.[77] This effect, however, is not a neoclassical price effect.[78] It produces relevant welfare effects on the corresponding parties that should be tracked in the distributional tax-reform table. It is possible that these non-standard effects persist in the long-run and are asymmetric for increases and decreases. The studies by Saez, Schoefer, and Seim (2019) for Sweden and Benzarti and Harju (2021) for Finland show that payroll tax incidence happens at the firm level rather than the individual level as in standard theory. Saez, Schoefer, and Seim (2019) show that firms which have many workers eligible for a specific payroll tax cut on young workers increase the wages of all their workers, not just the eligible workers, and that profits also go up. This suggests that, within the firm, workers and profits share the tax cut or tax increase in proportion to their share in value-added but there remains considerable uncertainty

[76]Saez et al. (2012) show that, in Greece, uncapping employer payroll tax increases the labor cost of the corresponding workers but uncapping the employee payroll tax does not. Bozio et al. (2025) find the same result in France when there is no close link between employer payroll taxes and benefits. Saez, Schoefer, and Seim (2019) find that employer payroll tax cuts for the young in Sweden do not increase their net-wages and businesses redistribute the tax cut windfall partly across all workers. Rubolino (2022) finds a female specific payroll tax cut does not increase their net-wages but boosts female employment and firm performance. Guillot (2019) shows that a special temporary employer payroll tax on very high wage earners in France was mostly borne by employers but then asymmetrically increased net wages upon expiration of the tax.

[77]Conceivably, it could also increase prices of output of the business benefitting consumers. There is no direct empirical evidence on this to date but the literature on minimum wage increases shows compellingly that part of this extra labor cost is passed on consumers (e.g. Harasztosi and Lindner 2019 find that 75% of the minimum wage).

[78]Bozio et al. (2025) provide a meta-analysis of 21 estimates in the literature showing that employer payroll tax changes are not passed to corresponding workers except when there a tight and salient tax-benefit linkage. While most of these studies tried to interpret their finds within the neo-classical supply vs. demand elasticities framework, Bozio et al. (2025) show that non-standard effects: saliency of the link between taxes and benefits and inequity aversion within firms are a more parsimonious way to account for the disparate empirical findings.

on how such findings generalize. It is likely that sharing depends on institution and bargaining power of workers within the firm as suggested by Kim, Kim, and Koh (2022).[79] Therefore, in Table A3, we tentatively assume that an employer payroll tax change would be borne collectively within the firm by workers for 2/3 and by profits for 1/3.[80] There is also evidence of strong employment effects of employer payroll tax changes particularly if tax changes are targeted to a specific group.[81] But with rigid wages, such employment effects fail to generate wage responses as predicted by standard incidence.

A.5.2 Application: Medicare for All Funding

The United States is the only advanced economy without universal health insurance. Nearly half of the population has to pay for their health insurance privately, primarily through employers. Employer-provided health insurance is part of the labor income of the corresponding workers. It has become a mandated benefit after the Affordable Care Act enacted in 2010 (except for small employers and part-time employees). Economically, compelling employers to provide insurance is similar to funding health insurance for workers with a payroll tax—but a very specific type of payroll tax: one equal to the cost of health insurance for the corresponding worker, as opposed to proportional to earnings as for usual payroll taxes. Saez and Zucman (2019), Case and Deaton (2020), and Finkelstein et al. (2023) analyze this regressive funding system—similar to a head tax per worker—and discuss how to create a more progressive form of financing.[82]

To illustrate the importance of non-standard tax incidence, consider replacing the current head-tax mandate by a payroll tax proportional to earnings for all workers currently covered by their employers. In 2021, about 85 million workers have employer-sponsored health insurance, covering about 155 million non-elderly individuals (see KFF, 2021). These workers have total pre-tax wage income of $9.0 trillion, their health insurance costs nearly $1.1 trillion, so the new payroll tax would need to be levied at a rate of about $\tau = 12\%$.[83] We assume that health

[79]They show that in Singapore, an employer payroll tax cut on workers aged 60 and above increases wages and they provide a meta-analysis of the literature suggesting that countries with more competitive labor markets show less evidence of non-standard incidence effects.

[80]Earlier macro-level studies have pointed out that the stability the labor share in national income in spite of large increases in employer payroll taxes in the 20th century suggests that profits are not affected by employer payroll taxes in the long-run (see e.g., Brittain 1971).

[81]See Saez, Schoefer, and Seim 2019 for youth in Sweden, Ku, Schoenberg, and Schreiner 2020 for local changes in Norway, Benzarti and Harju 2021 for small businesses in Finland, Cottet 2022 for low wage workers in France, Rubolino 2022 for female hires in Italy, Citino and Fenizia 2022 for apprentices in Italy.

[82]The current system is not exactly akin to a head tax as insurance cost varies by family size (if the policy also insures the family members of the worker) and by type of insurance provided. Moreover, although providing health insurance is mandatory for most employers, the premiums are not paid to the government but to private companies, and thus are best described as non-tax compulsory payments. We refer to them as a head tax to highlight the similarities with standard taxes. These payments are not included in our analysis of the progressivity of the current US tax system presented in Section 3 and Section 5.1 above.

[83]We assume that the payroll tax would be charged before any other tax and not be part of the tax base for existing payroll and income taxes, mimicking the current tax-exempt status of employer-provided health insurance.

insurance would remain the same worker by worker (to focus solely on the funding aspect—ignoring the complex issues of heterogeneity in benefits).

Conventional incidence with flexible wages. In the conventional analysis and in our neoclassical analysis, both the current head tax and the proportional payroll tax are taxes on labor, borne by the corresponding workers. This is also the assumption used in the recent comprehensive analysis of Finkelstein et al. (2023) who consider a shift to proportional payroll tax funding. This reform would leave labor costs (cash wage earnings plus the cost of health insurance and other fringe benefits) unchanged for each worker. After the reform, health insurance premiums currently paid by employers convert into extra gross cash earnings dollar for dollar and worker by worker. All gross earnings are then reduced proportionally by a factor $1-\tau$ due to the new payroll tax. Any worker with health care insurance costs above τ times her earnings benefits from the reform (and conversely). This is a progressive reform that would have significant positive effects on the disposable income of the working and middle class, at the expense of higher paid workers. Pre-tax incomes would not change, but post-tax incomes would become more equal, as the proportional payroll tax replaces the head tax for covered workers. As shown by Table A4, column 6, this reform would increase after-tax earnings of the bottom 50% by about 2.5% (and reduce after-tax income at the top by 1% to 2%), making it seemingly simple to put US health care funding on a fairer and more sustainable path.[84] Non-standard tax incidence, however, is crucial to understand how and whether this would work in practice. Three scenarios illustrate this point.

Employee payroll tax with rigid wages. Suppose first that the new payroll tax is charged to employees. Workers would see their net earnings reduced by the new payroll tax (as workers bear this new tax one for one, cf. Table A3). The elimination of the head tax is akin to an employer payroll tax cut. As explained in Table A3, because of wage rigidities, this payroll tax cut would not be passed one-to-one to the corresponding workers but instead passed roughly 2/3 to workers across the board within each firm (proportionately to their wages) and 1/3 to profits, according to existing studies. Under these assumptions, the reform becomes *regressive* as illustrated in Table A4. Both pre-tax and post-tax incomes become more unequal.

Employer payroll tax with rigid wages. Suppose now that the new payroll tax is charged to employers. The head tax–the current insurance premiums paid by employers–becomes an employer payroll tax. Because the amounts are the same, there is no tax savings or costs for

[84]The results in Table A4 report the effects of the reform on the full population including non-workers and on total pre-tax income (including non-wage income). Restricting the analysis to the covered population (less than half of the total population of the United States) and wage income only, the targeted population and income concepts would correspondingly show larger effects.

employers. Wage rigidities imply that net earnings do not change but pre-tax labor income becomes more unequal: labor costs for each worker change by the difference between the new payroll tax and the former head tax.[85] The reform again fails to make the progressive gains of the conventional analysis. But it is more progressive than the previous scenario as none of the savings made by employers goes to profits.[86]

With non-standard incidence effects of this kind, labor costs for workers change so that there could be employment effects due to labor demand responses of employers (cf. Table A3). In our setting, as labor costs for low-paid workers fall, they could become more attractive to employers, boosting employment at the low end. In a competitive standard labor market model, such demand effects lead to wage responses generating the conventional incidence results. But with rigid wages, such responses may be sluggish and incomplete, as shown by empirical evidence. There remains considerable uncertainty, and hence need for more research on how quickly such wage adjustments would take place.[87]

Directed tax incidence with rigid wages. Last, how can incidence be steered toward the equilibrium of the conventional analysis in the real world with wage rigidities? As proposed by Saez and Zucman (2019), existing employer-provided benefits could be converted one-for-one into a permanent wage increase worker-by-worker by law. This would leave labor costs for employers unchanged worker by worker. A new payroll tax on employees should then be created at rate τ as in the first scenario. This tax would fall on the corresponding employees. It would recreate the exact conventional incidence. The key difference with conventional incidence is that the equilibrium would be reached by legislation rather than through competitive market forces.

[85] This is the most plausible incidence in light of the studies analyzed above. Because there is a linkage between the head tax and health care benefits, the analysis of Bozio et al. (2025) suggests that the incidence passed on to workers individually (Gruber 1994 and Baicker and Chandra 2006 present US based analysis of health premiums changes consistent with this). Therefore, it is possible that wages would not be completely rigid and that the incidence of removing the head tax would eventually shift back to workers as in the standard incidence model.

[86] In the employee payroll tax scenario, if workers have a lot of power and can recoup 100% of the saving (instead of just 2/3), then it is likely that workers would insist on a proportional-to-earnings compensation to offset the payroll tax. In this case, the incidence is the same in the employee vs. employer payroll tax funding cases, but this equivalence depends crucially on strong worker bargaining power (instead of being the standard consequence of competitive markets).

[87] The classic study by Gruber (1994) on mandated maternity benefits in some US states found that wages of child bearing women adjusted within a few years.

Appendix Additional References

Auerbach, Alan J. 2019. "Tax equivalences and their implications." *Tax Policy and the Economy* 33, no. 1: 81-107.

Baicker, Katherine and Amitabh Chandra. 2006. "The Labor Market Effects of Rising Health Insurance Premiums," *Journal of Labor Economics* 24 (3), 609–634.

Benzarti, Youssef, and Dorian Carloni. 2019. "Who really benefits from consumption tax cuts? Evidence from a large VAT reform in France." *American Economic Journal: Economic Policy,* 11(1): 38–63.

Benzarti, Youssef, Dorian Carloni, Jarkko Harju, and Tuomas Kosonen. 2020. "What goes up may not come down: asymmetric incidence of value-added taxes." *Journal of Political Economy* 128, no. 12: 4438-4474.

Benzarti, Youssef, and Jarkko Harju. 2021. "Using payroll tax variation to unpack the black box of firm-level production." *Journal of the European Economic Association* 19, no. 5: 2737-2764.

Besley, Timothy J., and Harvey S. Rosen. 1999. "Sales taxes and prices: an empirical analysis." *National tax journal* 52, no. 2: 157-178.

Bozio, Antoine, Thomas Breda, Julien Grenet, and Arthur Guillouzouic 2025. "Does Tax-Benefit Linkage Matter for the Incidence of Payroll Taxes?" *Review of Economic Studies.*

Bradford, David F. 1986. *Untangling the Income Tax* Cambridge, MA: Harvard University Press.

Brittain, John A. 1971. "The Incidence of Social Security Payroll Taxes." *American Economic Review* 61(1), 110-125.

Carloni, Dorian, and Terry Dinan. 2021. "Distributional Effects of Reducing Carbon Dioxide Emissions With a Carbon Tax", Congressional Budget Office Working Paper 2021-11. No. 57399.

Case, Ann and Angus Deaton. 2020. *Deaths of Despair and the Future of Capitalism.* Princeton University Press.

Chancel, Lucas. 2022. "Global carbon inequality over 1990–2019." *Nature Sustainability* 5, no. 11: 931-938.

Chancel, Lucas and Yannic Rehm. 2023. "The Carbon Footprint of Capital: Evidence from France, Germany and the US based on Distributional Environmental Accounts," PSE Working Paper.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence." *American Economic Review,* 99(4): 1145–1177.

Citino, Luca and Alessandra Fenizia. 2022. "The Effectiveness of Targeted Payroll Tax Reductions", Working Paper.

Clemens Fuest, Andreas Peichl, and Sebastian Siegloch. 2018. "Do higher corporate

taxes reduce wages? Micro evidence from Germany." *American Economic Review* 108(2): 393–418.

Cottet, Sophie. 2022. "Payroll tax reductions for minimum wage workers: Relative labor cost or cash windfall effects?" Paris School of Economics Working Paper.

Eisinger, J., Ernsthausen, J., Kiel, P. 2021. "The secret IRS files: trove of never-before-seen records reveal how the wealthiest avoid income tax." *ProPublica*, June 8, 2021.

Finkelstein, Amy, Casey C. McQuillan, Owen M. Zidar, and Eric Zwick. 2023. "The Health Wedge and Labor Market Inequality." NBER Working Paper No. 31091, forthcoming *Brookings Papers on Economic Activity.*

Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *American economic review* 84(3), 622-641.

Guillot, Malka. 2021. "Tax Incidence among the Working Superrich. Evidence from the French 75% Tax on Millionaires. Evidence From the French 75% Tax on Millionaires", ETH Zurich Working Paper.

Hall, Robert E. and Alvin Rabushka. 1985. *The Flat Tax,* Stanford, CA: Hoover Institution Press.

Harju, Jarkko, Tuomas Kosonen, and Oskar Nordström Skans. 2018. "Firm types, price-setting strategies, and consumption-tax incidence." *Journal of Public Economics* 165: 48-72.

Kennedy, Patrick, Christine Dobridge, Paul Landefeld, and Jacob Mortenson. 2022. "The Efficiency-Equity Tradeoff of the Corporate Income Tax: Evidence from the Tax Cuts and Jobs Act."

KFF. 2021. "2021 Employer Health Benefit Survey," available online at `https://www.kff.org/report-section/ehbs-2021-summary-of-findings`

Kim, Jinyoung, Seonghoon Kim, and Kanghyock Koh. 2022. "Labor market institutions and the incidence of payroll taxation." *Journal of Public Economics* 209: 104646.

Kosonen, Tuomas. 2015. "More and cheaper haircuts after VAT cut? On the efficiency and incidence of service sector consumption taxes." *Journal of Public Economics* 131: 87-100.

Kotlikoff, Laurence J., and Lawrence H. Summers. 1987. "Tax Incidence." In *Handbook of public economics*, vol. 2, 1043-1092. Elsevier: North Holland.

Ku, Hyejin, Uta Schoenberg, Ragnhild C. Schreiner. 2020. "Do place-based tax incentives create jobs?" *Journal of Public Economics* 191, 2020, 104105,

Mankiw, Gregory. 2001. "Commentary: Balanced-Budget Restraint in Taxing Income From Wealth in the Ramsey Model." In *Inequality and Tax Policy* edited by K. A. Hassett and R. G. Hubbard. Washington, DC: AEI.

Piketty, Thomas. 2000. "Theories of persistent inequality and integenerational mobility." *Handbook of Income Distribution* volume 1, 429-476.

Piketty, Thomas, and Emmanuel Saez. 2013. "A Theory of Optimal Inheritance Taxation," *Econometrica* 81(5), 1851-1886.

Poterba, James M. 1996. "Retail price reactions to changes in state and local sales taxes." *National Tax Journal* 49, no. 2: 165-176.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *Economic journal* 37, no. 145: 47-61.

Rees-Jones, Alex, and Dmitry Taubinsky. 2020. "Measuring "Schmeduling"." *Review of Economic Studies* 87, no. 5: 2399-2438.

Rubolino, Enrico. 2022. "Taxing the Gender Gap: Labor Market Effects of a Payroll Tax Cut for Women in Italy", CESifo Working Paper No. 9671.

Saez, Emmanuel, Manos Matsaganis, and Panos Tsakloglou. 2012. "Earnings Determination and Taxes: Evidence from a Cohort-Based Payroll Tax Reform in Greece." *Quarterly Journal of Economics* 127 (1): 493–533.

Saez, Emmanuel, Benjamin Schoefer, and David Seim. 2019. "Payroll Taxes, Firm Behavior, and Rent Sharing: Evidence from a Young Workers' Tax Cut in Sweden." *American Economic Review.*

Slemrod, Joel. 1995. "Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986." *American Economic Review* 85(2), 175-80.

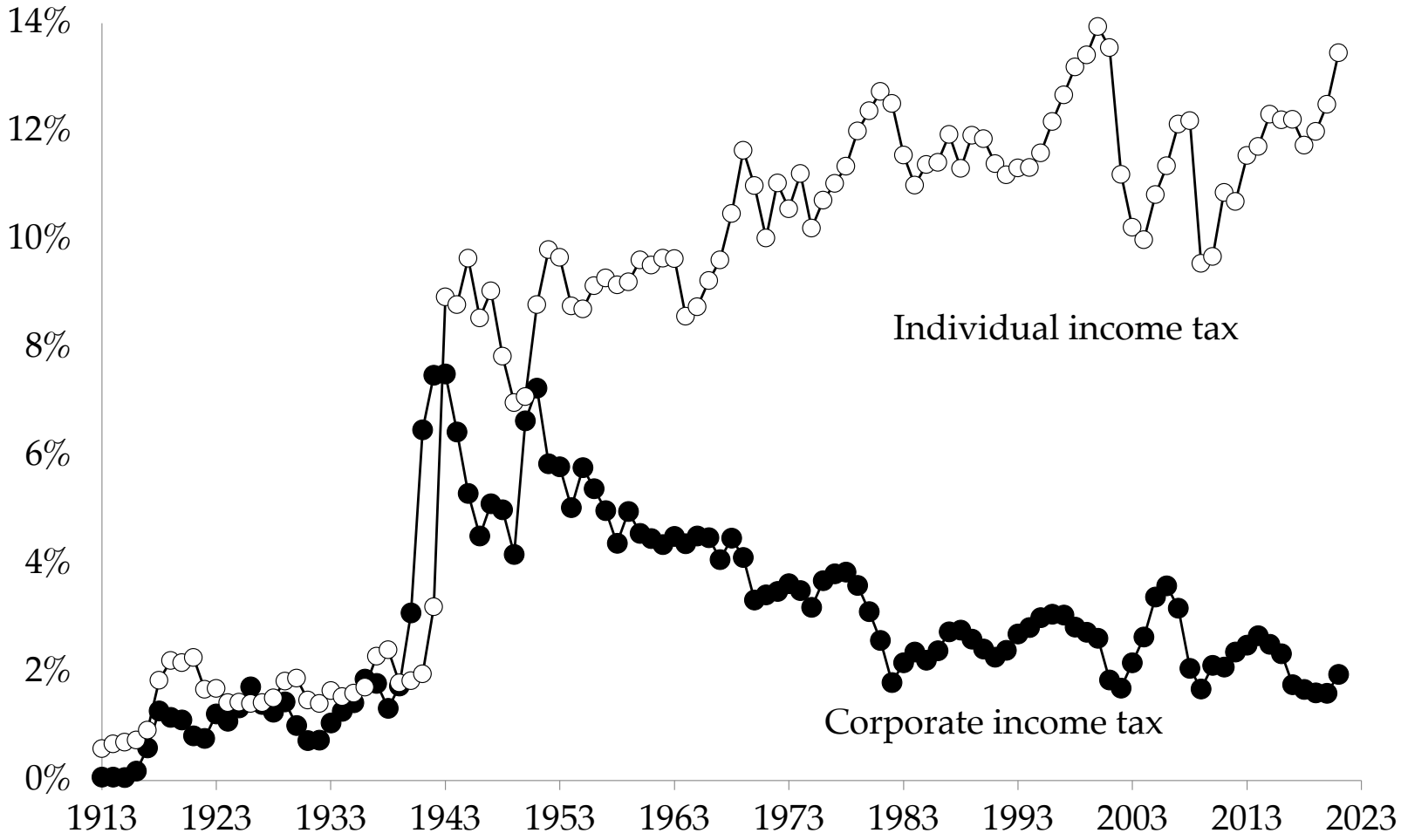
Stiglitz, Joseph E., and Partha Dasgupta. 1971. "Differential taxation, public goods, and economic efficiency." *Review of Economic Studies* 38, no. 2: 151-174.

Straub, Ludwig, and Ivan Werning. 2020. "Positive long-run capital taxation: Chamley-Judd revisited." *American Economic Review* 110, no. 1: 86-119.

Viard, Alan D., and Robert Carroll. 2012. *Progressive Consumption Taxation: The X-Tax Revisited.* AEI Press, 2012.

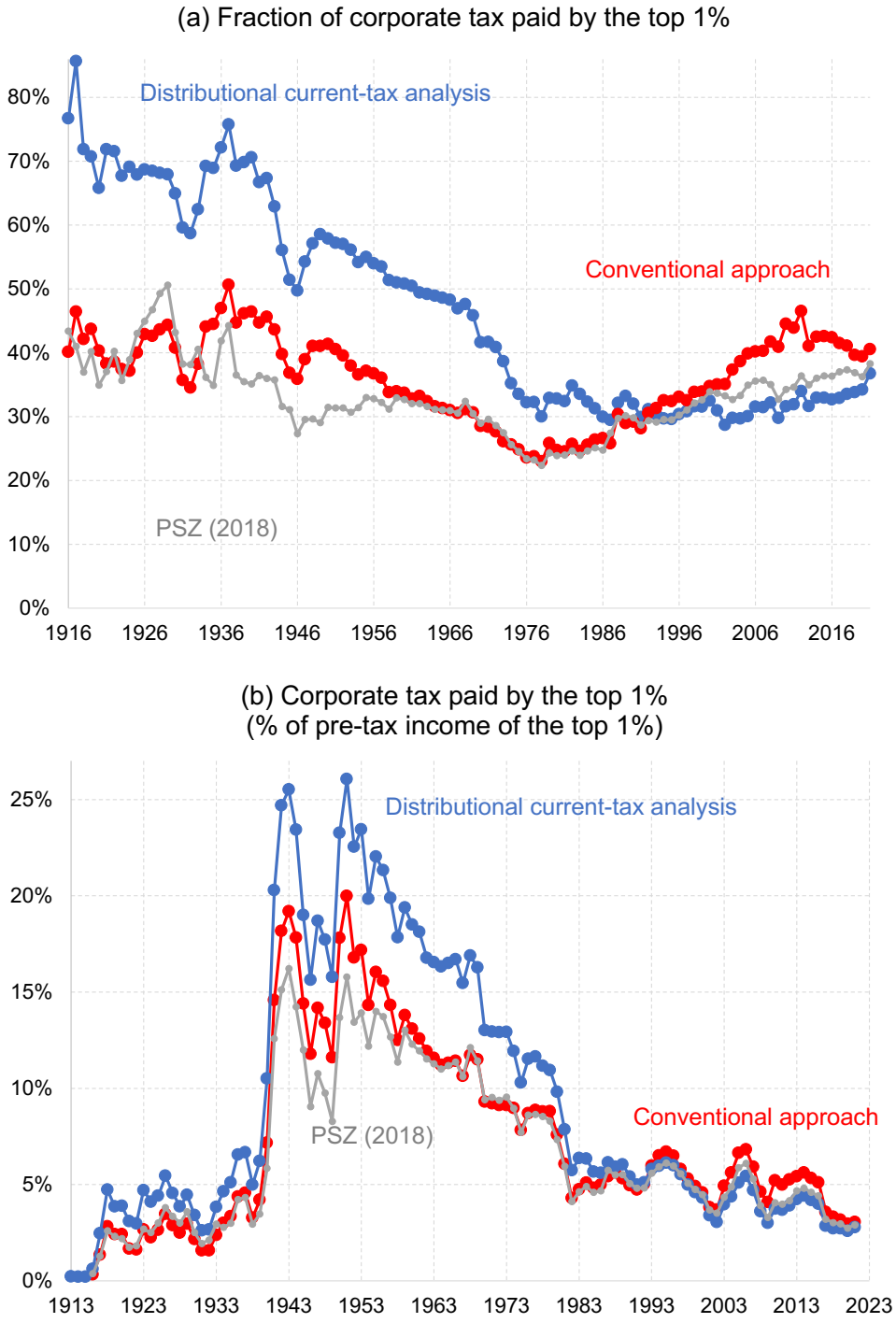
Zucman, Gabriel. 2023. "Cross-Border Fiscal Flows," work in progress.

Figure A1: Individual vs. Corporate Income Tax Revenue (% of National Income)



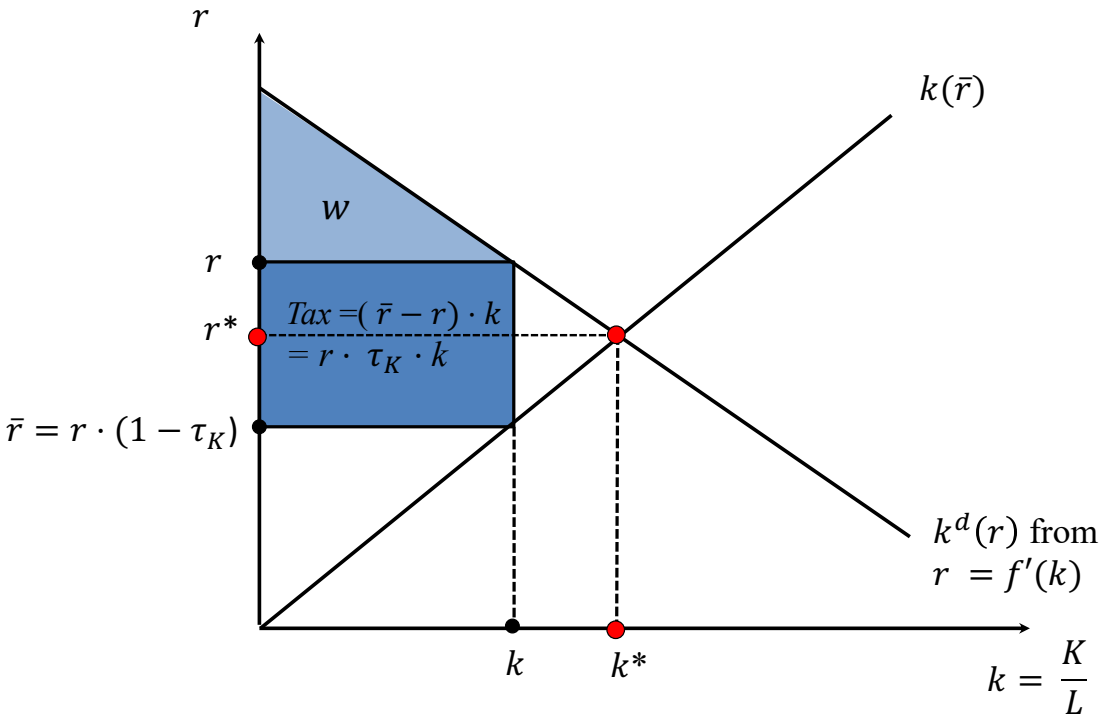
Notes: This graph shows the evolution of US corporate income tax revenues and individual income tax revenues from 1913 to 2021, expressed as a fraction of US national income. Federal, state and local taxes are included.

Figure A2: Allocating the Corporate Tax: Conventional Approach vs. Piketty-Saez-Zucman (2018) vs. Our Methodology



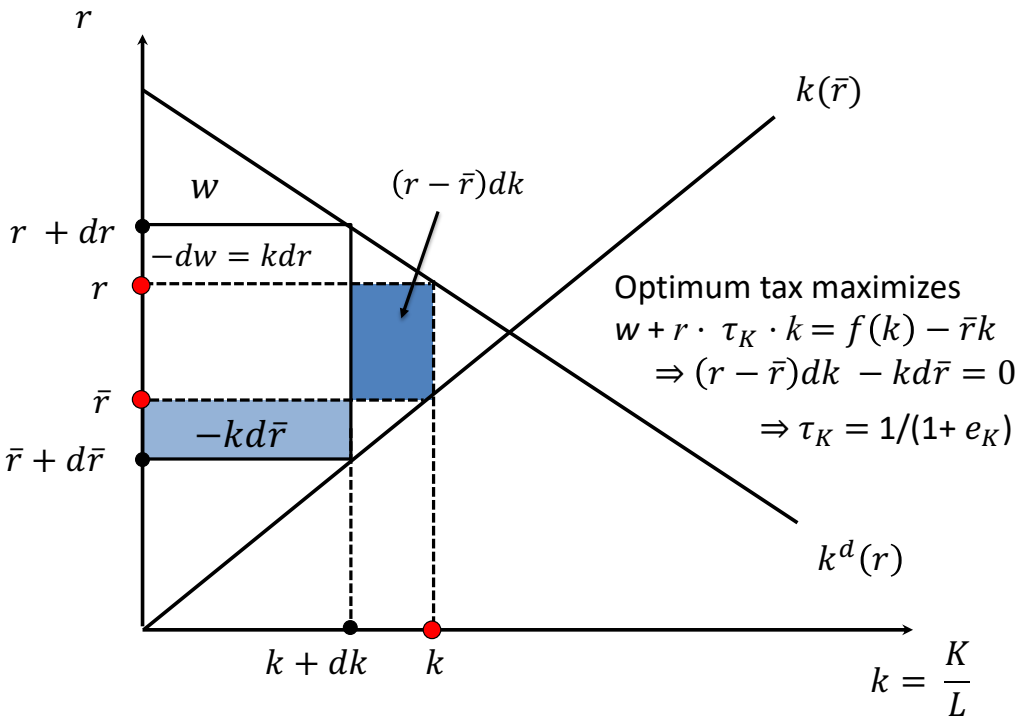
Notes: The top panel contrasts the share of the US corporate income tax (federal and state) paid by the top 1% units with the highest pre-tax national income in our current-tax methodology and the conventional practice of distributional tax analysis, as implemented by the Congressional Budget Office (CBO) and the original series of Piketty, Saez and Zucman (2018), denoted by PSZ. The bottom panel plots the amount of corporate taxes paid by the top 1% (as a fraction of the pre-tax income of the top 1%) implied by each of these methodologies.

Figure A3: General Equilibrium with Capital Tax



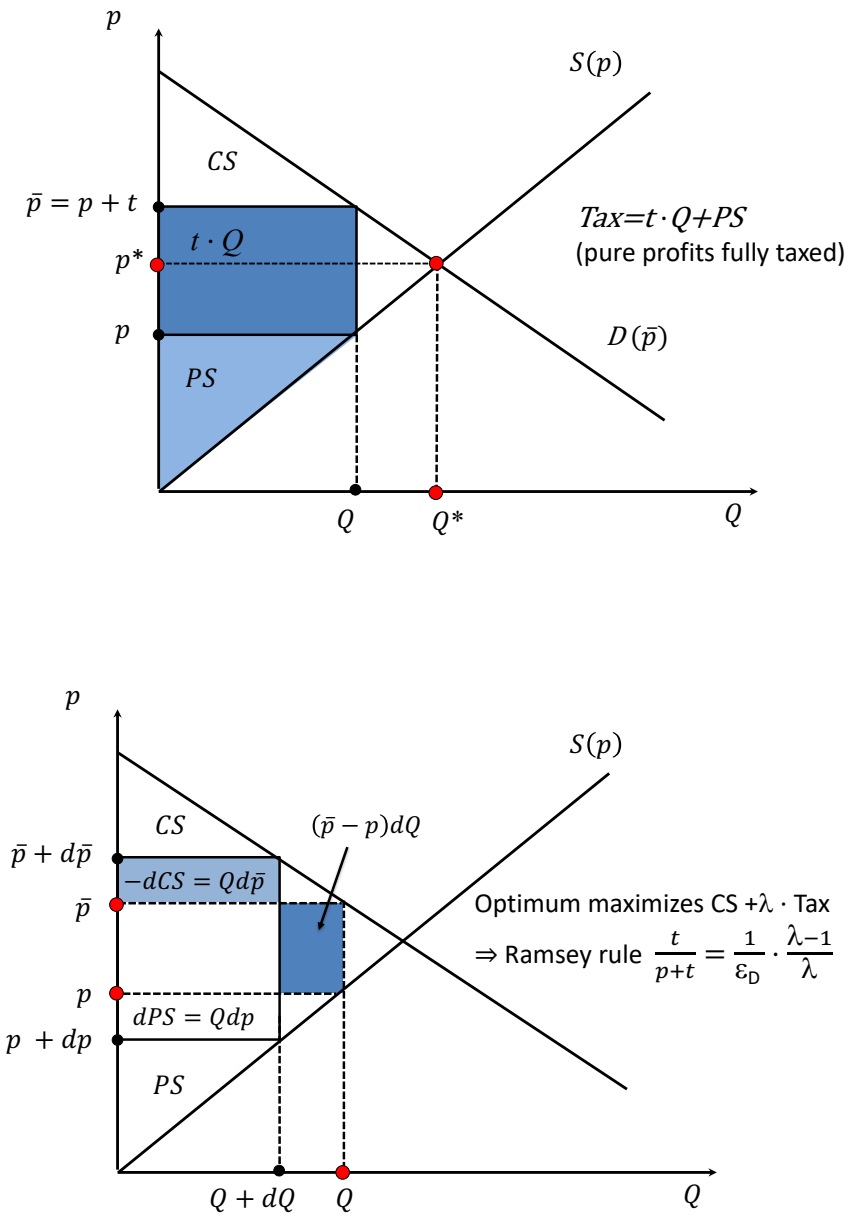
Notes: The figure depicts the effect of a tax on capital income at rate τ_K on the interest rate r, the capital to labor ratio $k = K/L$, and the wage w in a general equilibrium neoclassical model with fixed labor L, CRS production $F(K, L) = L \cdot F(K/L, 1) = L \cdot f(k)$. The equilibrium is characterized by 3 equations: (1) $r = f'(k)$ (rate of return of capital equals its marginal return which generates the demand for capital $k^d(r)$), (2) $k = k(\bar{r})$ (capital supply depends on its net of tax return $\bar{r} = r(1 - \tau_K)$), (3) $w = f(k) - kf'(k) = \int_0^k f'(\kappa)d\kappa - rk$ (the wage w can be read as the area below the demand curve and above the r horizontal line). Without taxes, the equilibrium is (r^*, k^*). With a tax rate τ_K, the equilibrium shifts to (r, k). The tax collects the rectangle, $(r - \bar{r})k = \tau_K rk$, it increases r, and reduces $\bar{r}$ and w. The tax reduces the wage and the surplus of capitalists by an excess burden triangle $\simeq (1/2) \cdot r\tau_K \cdot (k^* - k)$ over and above taxes collected. In this economy, pre-tax labor income is wL, pre-tax capital income is rK, and post-tax capital income is $r(1 - \tau_K)K$.

Figure A4: Capital Tax Reform and Optimum



Notes: The figure depicts the effect of a change $d\tau_K$ in the capital income tax rate τ_K in the simple neoclassical model depicted on Figure A3. The tax change reduces capital k by $dk < 0$, increases the pre-tax rate of return r by dr, reduces the net-of-tax rate $\bar{r}$ by $d\bar{r} < 0$. If the government wants to maximize the welfare of workers, it sets τ_K to maximize $w + \tau_K rk$ (wages plus tax revenue extracted from capitalists). As $w = f(k) - kf'(k)$, we have $w + \tau_K rk = f(k) - \bar{r}k$, the area below the demand curve $r = f'(k)$ and above the horizontal line $\bar{r}$ (the blue areas in Figure A3). The first order condition for the optimum is $(f'(k) - \bar{r})dk - kd\bar{r} = 0$ (the 2 blue rectangles on the Figure cancel out at the optimum). As $f'(k) = r$, this can be rewritten as $(r - \bar{r})dk/d\bar{r} = k$ or $(r - \bar{r})/\bar{r} = 1/e_K$ which is the classical inverse elasticity rule $\tau_K^* = 1/(1 + e_K)$ where $e_K = (\bar{r}/k)dk/d\bar{r}$ is the pure supply side elasticity. Therefore the classical pre-tax price incidence dr, dw is irrelevant for optimal tax analysis, a result that generalizes to any social welfare function as shown in Diamond and Mirrlees (1971).

Figure A5: Consumption Tax: Incidence and Ramsey Optimum



Notes: The top panel depicts the classic consumption tax incidence in a one good model. If we assume as in Diamond and Mirrlees (1971) that pure profits (=producer surplus in the diagram) can be fully taxed away, the tax is represented by the blue areas: $t \cdot Q + PS$. The bottom panel depicts the derivation of the optimum tax that maximizes consumer surplus plus taxes (weighted by factor $\lambda > 1$): $CS + \lambda T = v(Q) - \bar{p}Q + \lambda[\bar{p}Q - c(Q)]$. A small tax increase $d\bar{p}$ reduces CS by $Qd\bar{p}$ and increases taxes collected by $Qd\bar{p} + (\bar{p} - p)dQ$. Because pure profits are in the tax base, the increase in tax from the consumption good Qdp is fully offset by the loss of producer surplus dPS and hence the price effect dp is irrelevant. The first order condition $(\lambda - 1)Qd\bar{p} + \lambda dQ[\bar{p} - p] = 0$ leads to the classic inverse elasticity Ramsey rule $t/(p + t) = (1/\varepsilon_D) \cdot (\lambda - 1)/\lambda$. The supply side elasticity ε_S and the price effect dp are irrelevant.

Table A1: Current Federal Tax Distribution in the United States, 2021

	Pretax income		After-tax income		Taxes (federal only)		Tax rate composition (federal taxes only)					
Income groups	Average	Share	Average	Share	Share	Tax Rate	Individual income taxes	Payroll taxes	Consumption taxes	Property taxes (incl. estate tax)	Corporate tax	Corporate tax, conventional approach
P0-50	$20,889	12.3%	$17,862	13.1%	9.2%	14.5%	1.7%	10.5%	1.8%	0.0%	0.4%	0.8%
P50-90	$80,618	38.1%	$65,303	38.3%	37.3%	19.0%	6.8%	10.2%	1.0%	0.0%	1.0%	0.8%
P90-99	$243,587	25.9%	$195,098	25.7%	26.6%	19.9%	11.6%	6.2%	0.6%	0.0%	1.5%	1.3%
P99-99.9	$1,085,455	11.5%	$855,334	11.3%	12.6%	21.2%	16.5%	2.3%	0.4%	0.2%	1.8%	2.0%
top 0.1%	$10,288,542	12.2%	$7,956,531	11.7%	14.2%	22.7%	18.1%	0.8%	0.3%	1.1%	2.3%	2.9%
All	**$84,672**	**100%**	**$68,266**	**100%**	**100%**	**19.4%**	**9.9%**	**7.1%**	**0.8%**	**0.2%**	**1.3%**	**1.3%**

Notes: Groups based on pre-tax income including pure realized capital gains (defined as realized gains in excess of 3% of national income). Unit is individual adult (aged 20+) with equal split among couples. Pre-tax income is income before all taxes but after the operation of pension systems (public and private). Taxes include federal taxes only. Refundable tax credits are not included as negative tax (as they are treated as transfers, like other cash transfers, in the national accounts). Labor taxes assigned to corresponding workers, capital taxes to corresponding asset owners, consumption taxes to final consumers. In the conventional approach (currently used by CBO), the corporate tax is assigned 75% to capital income on individual tax returns and 25% to labor income (with no adjustment for corporate profits earned through pension funds).

Table A2: Illustration of Current-Tax analysis: Case Studies (2018)

Millions of US$	Jeff Bezos	Warren Buffett
US federal taxes	**43**	**930**
Individual income tax	43	5
Corporate tax	0	925
Payroll taxes	0	0
Consumption taxes	0	0
US state and local income taxes	**140**	**241**
Individual income tax	0	1
Corporate taxes	70	53
Business property taxes	69	187
Consumption taxes	~0	~0
Residential preoperty taxes	~0	~0
Foreign taxes	**154**	**337**
Corporate taxes	123	337
Business property taxes	31	0
Total taxes	**337**	**1,508**
Pre-tax income	2,221	8,176
Effective tax rate	**15.2%**	**18.4%**
Federal	1.9%	11.4%
State and local	6.3%	2.9%
Foreign	6.9%	4.1%

Notes: See text for complete sources and details. Corporate taxes paid are equal to global cash tax payments reported by Amazon and Berkshire Hathaway in their SEC 10-K reports, apportioned by the ownership stake of Bezos and Buffett respectively. No geographical breakdown of cash taxes paid is available. We use the published breakdown of provisions for current taxes (Amazon) and provisions for current plus deferred taxes (Berkshire Hathaway) to allocate these cash payments to federal vs. state and local vs. foreign governments. Property taxes are computed as 1% of net property and equipment, and allocated to US state and local governments vs. foreign governments based on the geographical location of assets reported in the 10-K item 2. Individual income taxes are taken from Eisinger et al. (2021) for federal taxes and based on public information about state of residency for state and local taxe. State and local consumption and residential property taxes are assumed to be negligible relative to income. Income is equal to the apportioned share of Amazon and Berkshire Hathaway's pre-tax profits (excluding unrealized gains on investments, and adding imputed business property taxes) plus any individual income (e.g., realized capital gains, wages, income from other investments) identified in Eisinger et al. (2021).

Table A3: Lessons from the Modern Literature on Non-Standard Tax Incidence

Tax	Who bears the burden of a tax change	Notes and key references	Nature/hierarchy of main behavioral Responses	Size of behavioral Responses
	(1)	(2)	(3)	(4)
Individual income Tax	Individuals 100%	Consistent with conventional incidence	Avoidance/evasion Real responses	Varies with context, can be large Likely small. Inattentiveness (Rees-Jones, Taubinsky 2020)
Corporate income tax	Profits 2/3* Workers 1/3* Consumers 0%*	Fuest, Peichl, and Siegloch (2018) for Germany and Kennedy et al. (2022) for the US. Likely depends on bargaining power. Asymmetric effects?	Avoidance/evasion Real responses	Varies with context, can be large Likely medium, varies with design
Consumption taxes				
Value-added-tax or excise tax **increase**	Consumers 100%	Benzarti et al. (2020) on VAT in Europe	Evasion Consumer demand	Varies with context, can be large Larger response for tax on specific goods
Value-added-tax or excise tax **decrease**	Consumers 50% Profits 37.5%* Workers 12.5%*	Benzarti et al. (2020) on VAT in Europe Benzarti and Carloni (2019). Likely depends on bargaining power	Consumer demand	Response muted by 50% price passthrough
Sales taxes (not posted on prices)	Consumers 100%	Consistent with conventional incidence. Poterba (1996) and Besley and Rosen (1999) for local sales tax in the US	Evasion Consumer demand response	Can be large for small retailers Muted by inattentiveness (Chetty et al. 2009)
Payroll taxes				
Employee side payroll tax	Workers 100%	Consistent with conventional incidence	Labor supply response	Likely small (higher for less attached subgroups)
Employer side payroll tax	Corresponding workers 0%	Saez et al. (2012) for Greece, Bozio et al. (2022) for France, Saez et al. (2019) for Sweden	Employer labor demand responses	Can be large for targeted tax changes
	Workers **collectively** 2/3* Profits 1/3* Consumers 0%*	Saez et al. (2019) for Sweden, Benzarti and Harju (2021) for Finland. Likely depends on bargaining power. Asymmetric effects?		

Notes: Column 1 reports who bears the burden of a tax change with some explanatory notes and key references in column 2. A * denotes large uncertainty in the estimate, and therefore where further research would be most valuable. The table ignores any neoclassical pre-tax price effects as they are normatively irrelevant and hard to compelling estimate empirically. Therefore, incidence is always within a production unit (such as a firm) on how surplus is shared among stakeholders in the unit (owners profits, workers earnings, and consumers' prices). Column 3 lists the most important behavioral responses with some notes on magnitudes in col. 4. "Small" means elasticity of the tax base with respect to the net-of-tax rate in range (0,.25), "medium" in range (.25,.5), "large" is .5 or more. See text for more detailed justifications and more nuanced explanations.

Table A4: Replacing Employer-Provided Health Insurance Contributions By a Payroll Tax

Income groups	Current system			Reform replacing current employer health care contributions by flat 11.8% payroll tax								
				Conventional incidence and directed incidence			Employee payroll tax with rigid wages			Employer payroll tax with rigid wages		
	Average pre-tax income	Current head tax ($ per adult)	Current head tax (% pre-tax income)	New payroll tax (% pre-tax income)	% change in pre-tax income	Change in after-tax income (% pre-tax income)	New payroll tax (% pre-tax income)	% change in pre-tax income	Change in after-tax income (% pre-tax income)	New payroll tax (% pre-tax income)	% change in pre-tax income	Change in after-tax income (% pre-tax income)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
P0-50	$20,889	$1,440	6.9%	4.5%	0.0%	2.4%	4.5%	-3.3%	-0.9%	4.5%	-2.4%	0.0%
P50-90	$80,618	$6,505	8.1%	7.0%	0.0%	1.1%	7.0%	-2.1%	-1.0%	7.0%	-1.1%	0.0%
P90-99	$243,587	$7,826	3.2%	5.2%	0.0%	-1.9%	5.2%	2.1%	0.2%	5.2%	1.9%	0.0%
P99-99.9	$1,085,455	$6,212	0.6%	2.7%	0.0%	-2.1%	2.7%	3.5%	1.4%	2.7%	2.1%	0.0%
top 0.1%	$10,288,542	$5,841	0.1%	1.3%	0.0%	-1.3%	1.3%	3.8%	2.5%	1.3%	1.3%	0.0%
All	**$84,672**	**$4,259**	**5.0%**	**5.0%**	**0.0%**	**0.0%**	**5.0%**	**0.0%**	**0.0%**	**5.0%**	**0.0%**	**0.0%**

Notes: This table simulates the distributional effects of replacing the premiums paid by employers for health insurance provided to their workers by a flat payroll tax in 2021. The total amount of employer-provided health insurance premiums is taken from the National Health Expenditures accounts, Table 5.6, sum of contributions to employer-sponsored private health insurance paid by private business, households, federal government, and state and local governments. The total amount is $1,068 billion in 2021, which is equal to 5.0% of total national income (including pure realized capital gains) and 11.8% of the total pre-tax wage income of currently-covered employees. This total is allocated to income groups following the distribution of health insurance contributions reported in W2 forms (with a correction at the bottom of the distribution to take into account that only employers with more than 250 workers have to report). In column 3, the result is divided by pre-tax national income (as reported in Table 1 and in col. 1 here) to compute the current "head tax" rate. Columns 4 to 12 consider the effects of replacing this head tax by a flat payroll tax of 11.8% on the gross wage earnings of currently-covered employees. In cols. 4 to 6 we assume that health insurance premiums currently paid by employers convert into extra gross cash earnings dollar for dollar and worker by worker, so that pre-tax income does not change, and after-tax incomes rise at the bottom of the distribution and fall at the top (as a head tax is replaced by a flat tax). In cols. 7 to 9 we assume that the payroll tax is charged to employees, wages are rigid, 2/3 of what was previously paid by employers to insurers goes to covered workers and 1/3 goes to profits. In this case the reform is regressive: both pre-tax and after-tax income become more unequally distributed that in the current status-quo. In cols. 10 to 12 we assume that the tax is charged to employers and wages are rigid, so that pre-tax income increases by the difference between the payroll tax and the head tax (col. 10 minus col. 3), and there is no change in after-tax income.