Dynamic Analysis, Welfare, and Implications for Tax Reform

Jason Furman
Chairman, Council of Economic Advisers

National Bureau of Economic Research Tax Policy and the Economy Conference
Washington, DC
September 22, 2016

This is an expanded version of these remarks as prepared for delivery.

I work on a wide range of economic policies. While each area raises unique issues, the ultimate question in analyzing any of them is the same: what impact will this policy have on welfare? In many of the topics I work on, there is not a well-developed, coherent framework for answering that question. As a result, we often have to rely on heuristics: what stimulus measure has the most bang-for-the-buck? What training program results in the highest earnings? What intellectual property regime increases the quality of patents? In this regard, all of you working in public finance are lucky: you have a well-specified, coherent framework for assessing the welfare impacts of changes in tax policy, a framework that many of you here have helped to build and improve.

But while you are fortunate insofar as you have a ready-made framework at your disposal, you are rather unfortunate in that many policymakers discuss tax policy in ways that can depart significantly from what we know from our welfare frameworks. And the substance and presentation of the analysis of tax policies have sometimes played a role in furthering that confusion rather than helping to dispel it.

In my remarks today, I will try to offer some constructive criticism about the ways in which that analysis has been presented and used, as well as some sense of how it might be done differently. However, I should state at the outset my tremendous respect for what economists have accomplished in this area, how clear-thinking everyone in this room is about these issues, and how challenging it is to find simple metrics that will satisfy the interests and questions that policymakers ask.

My remarks today are not about dynamic scoring for official budgetary purposes, but instead are focused on the substance of dynamic *analysis*, its presentation, and its implication for evaluating alternative tax reforms. Specifically, I will make three arguments. First, growth is not the same as welfare, and evaluating policies solely through their effects on growth (or through their macroeconomic feedback on the budget deficit) implicitly adopts an unreasonable welfare criterion. Second, this is not just a minor technical point but fundamental: many tax reforms generate their growth effects through reductions in the consumption of non-market goods (e.g., leisure) or through tax increases on more moderate-income workers. As a result, the third point is that when analyzing a change in tax policy, the traditional, static distribution table is often the most useful information for evaluating its welfare impacts.

I should also state at the outset that economic growth is critically important for our economy, and we should adopt a wide range of policies to boost our growth rate. However, we need to distinguish between two phenomena: (i) the sustained increases in productivity that are the engine of long-run growth and that have sharply increased living standards over periods of multiple decades or more; and (ii) the ultimately constrained growth generated by, for example, people working longer hours in response to policy changes. The focus of public policy should be on the former, and I believe tax policy can play a part in promoting this type of growth. But in evaluating alternative tax policies, and especially in evaluating tax policies that entail significant tradeoffs between distributional progressivity and more intensive use of inputs, I think analytical rigor is essential to ensure that we do not conflate these two types of growth and thus ignore their sharply different implications for social welfare.

The Contrast Between Regulatory Analysis and Dynamic Analysis

Before getting to tax policy, I want to start with regulatory policy, which is one area where practitioners do welfare analysis in a conceptually correct way when deciding between alternative policies. Consider, for example, a decision about whether or not to promulgate a regulation that would require factories to install a piece of equipment that would prevent local pollution. One approach to this decision is cost-benefit analysis: do the benefits of the policy change (in this case, the number of lives saved by reduced pollution multiplied by the value of a statistical life) exceed its costs (in this case, the cost of installing the equipment multiplied by the number of factories that would have to install it)? An alternative approach would be to judge the policy based on the net impact of the regulation on jobs (or output), including the additional jobs (or output) created installing the pollution-reducing equipment less the jobs (or output) lost as a result of the higher production costs.

Just about anyone who has studied public policy could tell you that of these two options, the cost-benefit framework is the right one to answer this question and the jobs framework is the wrong one. In this case, cost-benefit analysis provides a coherent basis for a welfare calculation about the impact of diverting resources from other uses towards pollution control. In contrast, the jobs analysis is confused. For example, the jobs installing the pollution-reducing equipment should not be viewed as a net plus for society because these are jobs diverted from other productive activities that also have value.

Federal agencies considering regulations, and the Office of Information and Regulatory Affairs (OIRA), put the right items on the cost and benefit side of the ledger in making regulatory decisions. This is not to say this framework provides a definitive answer; there will always be uncertainty and debate about both basic parameters (e.g., the value of a statistical life) and the impacts of any given policy (e.g., the long-run costs of the pollution-reducing equipment or number of lives saved). But at least a decent amount of both the public debate and court review of regulations genuinely does focus on debating these parameters within a coherent welfare framework.

Dynamic Analysis is Not Welfare Analysis: Efficiency Issues

In contrast, the way in which policymakers interpret dynamic analysis can foster confusion rather than clarity in evaluating alternative tax policies. In particular, a singular focus on the growth effects (or even worse, the purported effects on jobs) of tax policy can be misleading. Finding that a given policy increases output by one percent is not the same as finding that a policy makes everyone one percent better off. In fact, the welfare impacts of a policy change can be—and often are—substantially smaller or larger than its impact on growth. (Note, throughout this discussion I use output and growth more or less interchangeably—after all, even a one-time increase in the level of output will result in a higher growth rate for a period of time.)

To simplify, consider a proposal to reduce the labor tax rate from 30 percent to 20 percent that leads someone to work an additional five hours a week at a job that pays $20 an hour (see Table 1). A dynamic analysis would find that this adds $100 per week to the size of the economy. But this analysis misses that those five hours were valuable to the person, and that as a result of the longer hours of work she now has less time, say, with her children or at the movies. Moreover, we know just about how much she values these activities. This knowledge is not based on our judgment of the value of time spent with children or at the movies, but instead on her revealed preferences. In particular, when her after-tax wage was $14 she thought that, at the margin, an hour of leisure was better than an hour of work, and when her after-tax wage was $16 she thought the opposite—putting the value of leisure somewhere between the two at, say, $15 an hour.[1] This observation provides the inputs for a coherent welfare analysis: the benefits of the policy change are the $100 in additional output, the costs are the $75 in foregone leisure ($15 per hour multiplied by five hours), and the net benefits are $25—which is one-quarter of what the output-oriented perspective would have told you.[2]

[1] Standard economics treats work as a cost in utility functions. This is likely the right treatment for the intensive margin where one is considering trading off the market consumption increase derived from additional hours of work against foregone leisure. But it may be wrong for the extensive margin, the choice between working and not working, when in many cases people derive substantial benefits from the dignity associated with working.

[2] In equation form using standard notation, an optimizing worker satisfies $u_c(1-t)w + u_l = 0$, and the welfare impact of changing the tax rate is $\frac{\partial V}{\partial t} = u_c w \frac{\partial l}{\partial t} + u_l \frac{\partial l}{\partial t} = \left(u_c(1-t)w \frac{\partial l}{\partial t} + u_l \frac{\partial l}{\partial t}\right) + u_c t w \frac{\partial l}{\partial t} = u_c t w \frac{\partial l}{\partial t}$, where the last equality relies on the worker's optimization condition. If the tax rate is around 25 percent, the welfare impact will be 25 percent of the increase output multiplied by the marginal utility of consumption.

Table 1
Illustrative Welfare Analysis of a Reduction in Labor Taxes
(Assuming a Representative Worker to Whom Tax Revenues Are Rebated Lump Sum)

	Baseline (30% Labor Tax)	Alternative (20% Labor Tax)
Social Benefit (Change in Output)		
Hourly Wage	20	20
Hours Worked per Week	40	45
Output per Week	800	900
Change in Output		**100**
Social Cost (Change in Value of Leisure)		
Hourly Value of Leisure	15	15
Leisure Hours per Week	80	75
Value of Leisure per Week	1,200	1,125
Change in Value of Leisure		**-75**
Net Social Benefits		**25**

Source: CEA calculations.

While this a special example, the point holds more generally, as does the rough quantification that, for many classes of tax policies, the welfare effects about are about one-quarter the size of the output effects. This is because many of the policies that are designed to raise output do so by inducing people to substitute market consumption for leisure consumption (i.e., to work more) or future consumption for present consumption (i.e., to save more, leading to more capital accumulation). However, assuming that each worker was making the best choice for herself given her situation before the change in tax policy, the changes in behavior resulting from that policy change have little direct value. This observation is simply a reflection of the well-known fact that small changes relative to an optimum have only a second-order impact on welfare because the utility curve is essentially flat near its peak. The larger societal gains come from the impact of her decisions on the government budget balance which can be a first order effect because *ex ante* government policy was distortionary and thus, unlike individual behavior, was not at a local maximum. It is for this reason that the social value of the tax change is roughly 25 percent of the increase in income, since 25 percent is a reasonable proxy for the actual marginal tax rate.

There may be cases in which the welfare effects of a tax reform come closer to its effect on growth. A change in tax policy that raised the level of output by raising Total Factor Productivity (TFP) without reducing leisure or current consumption, like expanding the Research and Experimentation credit to help firms internalize more of the spillover benefits of their research,

could be one such example. This illustrates that understanding the sources of the increased GDP resulting from a tax change is important. Indeed, it is likely that a tax reform that adds 0.5 percentage point to GDP through higher TFP is better from a welfare perspective than one that adds 1 percent to GDP through additional hours of work.

On the other hand, many tax policies have welfare effects that greatly exceed their output effects, including policies that have positive welfare effects in the absence of output effects. Consider, for example, the excise tax on high-cost employer-sponsored coverage, sometimes called the Cadillac tax. Employer contributions to health insurance are excluded from the income tax base for both workers and firms. Beginning in 2020, the high-cost excise tax will impose a 40 percent tax on any plan costs in excess of a projected $10,700 for individual coverage and $29,000 for family coverage, a provision affecting 7 percent of participants and 1 percent of plan dollars. Based on estimates from the Congressional Research Service, the Congressional Budget Office, and the Joint Committee on Taxation, the excise tax is estimated to result in a $40 billion reduction in health expenditures in 2024 and a $50 billion increase in wages in 2026 (Furman and Fiedler 2016). Even absent any effect on output, the improved composition of consumption would constitute a welfare improvement, specifically the $50 billion in wage increases net of the value of the foregone health spending, since previously individuals were overspending on relatively lower value health insurance and underspending on other goods and services. Using the same logic as the above, the foregone health expenditures will be worth about one quarter less than the increased consumption of other goods and services the tax makes possible. In practice, it is plausible that the value of these health expenditures is considerably lower if there are substantial inefficiencies in the health care sector, such as misaligned innovation incentives.

Again, the high-cost excise tax is not just a special case. Many of the largest potential benefits of individual tax reform result not from the traditional reform paradigm of a reduction in marginal rates leading to a greater incentive to work (more on this below), but from reforms that improve the allocation of consumption by reducing differential tax rates on different forms of consumption.

Dynamic Analysis is Not Welfare Analysis: Distributional Issues

So far I have considered efficiency issues alone. But another issue with relying on output effects is that doing so implicitly adopts an arbitrary and—for most people—undesirable social welfare function to judge the impact of policies. In particular, analyses focused solely on output increases assume that aggregate welfare is a linear function of total dollars earned—so that society would be made better off by a policy that decreased a poor person's income by just under $10 in order to increase the income of a rich person by $10. Alternatively, another social welfare function is that society would be made better off by a policy that decreased a poor person's income by just under 10 percent in order to increase the income of a rich person by 10 percent. This second social welfare function says we should look at changes in the average of the logs rather than changes in the average of the levels—but even that tradeoff would strike many of us as unattractive.

To give another toy example, consider a policy that increased the income of a high-income person and lowered it for two lower-income people as illustrated in Table 2. The policy would increase mean income—analogous to the concept of growth—by 4.7 percent. But the policy would reduce the mean of log income by 0.7 percent. By the first social welfare function, then, this would be a beneficial policy; but by the second, we should reject the proposed policy for decreasing social welfare on net.

Table 2
Illustrative Welfare Analysis of Changes in the Distribution of Income

	Baseline	Alternative	**Percent Change**
Individual Incomes			
Person 1	100	85	**-15.0**
Person 2	200	195	**-2.5**
Person 3	10,000	10,500	**5.0**
Aggregate Welfare			
Mean Income	3,433	3,593	**4.7**
Mean Log Income	6.37	6.32	**-0.7**
Mean CRRA Transformation of Income with $\theta = 2$[1]	0.99	0.99	**-0.1**

Source: CEA calculations.

[1] CRRA refers to the constant relative risk aversion functional form: (y^(1-θ)-1)/(1-θ).

Of course, the utilitarian framework shown here is not itself without controversy (e.g. Mankiw 2013). More generally, we do not have any undisputed social welfare function, and the preferred function varies greatly between people and among different political perspectives. Using growth as a metric, however, does not sidestep this debate; it just takes a particular side, one that is at variance with the way just about everyone appears to think about tax policy. After all, even the most conservative tax plans levy taxes proportional to income or consumption and not in a lump-sum manner. In other words, such plans assume that abilities to pay should be compared on a percentage basis (as in the log social welfare function) not an absolute dollars sense (as in the growth social welfare function). In fact, the log social welfare function likely understates society's views on distribution—considering that even the classic flat tax proposal actually had average rates rising as a function of income (Hall and Rabushka 1995). For many people, a social welfare function that is consistent with their philosophical views—and perhaps even the reveled preference of society in many circumstances—would place greater weight on percentage changes for lower-income households than higher-income households. For example, social welfare might be better captured by a Constant Relative Risk Aversion (CRRA) transformation of income with the parameter for relative risk aversion $\theta = 2$, as shown in the last row of Table 2.

An Illustrative Application: Substituting Lump Sum Taxes for Proportional Labor Taxes in a Ramsey Model

In many areas of public policy, the example shown in Table 2 would be a theoretical curiosity but would not meaningfully alter our ranking of different policy options. This is because, for many policies, we do not have any good way to determine the impact of each option on the distribution of income, and the assumption that it will have a uniform impact is not unreasonable. For example, growth may be a reasonable way to decide between different allocations of infrastructure or research funding.

But in tax policy, we often *can* estimate the static impact that a policy has on the distribution of income reasonably precisely, and this estimate is already part of the models we are using for revenue estimates and dynamic analysis. Moreover, these static distributional effects are often a first-order contributor to the economic effects—so considering one without incorporating the other can be deeply misleading.

To consider another example, I will analyze a 10-percent reduction in labor taxes paid for by a lump-sum tax. In a standard economic model, replacing distortionary taxes with non-distortionary taxes would increase the level of output. In fact, using the Ramsey framework put forward by N. Gregory Mankiw and Matthew Weinzierl (2006), as well as the particular parameter values selected by the authors, this policy would increase output by 1 percent and raise welfare by an amount equivalent to a 0.5-percent increase in consumption in the long run for the representative agent. As discussed above, the welfare effect is substantially less than the growth effect, but if aggregate output and the utility of a theoretical representative agent were the only considerations, the positive sign on both would be an indication that the policy would be worth pursuing.[3] (As an aside, I am using this model as an illustration and not endorsing any particular model or parameters for dynamic analysis. In particular, the assumption reflected in this model that capital is supplied perfectly elastically is important to the magnitude of the estimated effects on output for certain classes of tax reforms and unlikely to be a good approximation for the U.S. economy.)

However, the additional growth associated with the tax reform in this model arises precisely because the policy is making the tax code considerably less progressive. And in a world with more than one type of person, this reduction in progressivity matters for welfare. Assuming the income levels in the 2010 IRS public use file and the linear tax system of the Mankiw and Weinzierl article, the static cost of a 10-percent reduction in labor income taxes would be about

[3] Part of the reason the welfare effect is larger than the one-quarter rule of thumb I stated earlier is that this comparison is between utility levels in long-run equilibrium and thus overstates the welfare increase attributable to the tax reform, which would need to take into account that the benefits are smaller along the transition path between the initial steady state and the new steady state.

$150 billion.[4,5] Applying the percentage changes in economic variables derived from the macroeconomic model to the reported income values for each family in the public use file, the dynamic revenue feedback would offset about 12 percent of this cost and financing the remainder would require a lump-sum tax of about $900 per family.[6] For households making up to about $37,000 in labor income per year, this lump-sum tax would be larger than the gross savings from the 10 percent rate reduction. In fact, while all working families would see their before-tax incomes go up in this model as a result of their choice to work more, roughly two-thirds of families would also see their taxes go up solely as a result of the shift to lump-sum taxation (before accounting for behavioral changes). Note this analysis assumes that the *ex-ante* tax system was a proportional 25 percent rate with no exemptions; the number of losers from a proportional reduction in the actual progressive U.S. income tax cut financed by a shift to lump sum taxation would be considerably greater.

Table 3 presents a dynamic distributional analysis that reports the net effects of the labor income tax cut and the lump-sum tax on after-tax income and on utility. Unlike the examples discussed earlier, the utility impact is no longer proportionate to the change in output. While the increase in output resulting from the reduction in distortionary taxation remains the same, that increase in output results from a substantial redistribution from poor families to rich families, and that redistribution reduces utility for the majority of families. For example, consider households in the second quintile. On average they would see a $570 net increase in their taxes from the shift to lump sum taxation—which would not be offset by the additional $180 they earned working more hours. Households in the middle quintile would roughly break even with earnings up $310 and taxes up $340, but with hours up 1 percent, the overall impact on utility would be equivalent to a 0.6 percent reduction in consumption. Families in the top two quintiles, however, would be better off on average.

[4] Labor income is defined as the sum of wages and self-employment income. Self-employment income is limited to the Social Security taxable maximum in the public use file and this limitation is carried over into the current analysis. Filers with no wage income, no (or negative) capital income, and no transfer income are excluded. Filers with capital losses exceeding $10,000 are also excluded. Capital income is set to zero for all other filers with capital losses. The analysis is limited to families that file a tax return. As most families that do not file a return have little income, excluding them tends to bias the analysis in the direction of overstating the welfare gains from the tax cut.
[5] The $150 billion static cost does not include any behavioral responses to the tax reform whatsoever. It is not comparable to a conventional estimate produced by the Joint Committee on Taxation or the Department of the Treasury, which would include microeconomic behavioral responses such as changes in avoidance behavior.
[6] The 12 percent offset reported here differs from the offsets reported in the Mankiw and Weinzierl article for two reasons: (i) the ratio of capital income to labor income in the data underlying this exercise does not match that implied by the economic model used in the paper and (ii) the offset is computed for a discrete change in the tax rate not a differential change. To avoid the issue of extreme utility losses for families for which the lump-sum tax reduces income to (or below) zero, the lump-sum tax is reduced in such cases so as not to reduce after-tax incomes below $1. The tentative lump-sum tax amount is adjusted to raise the necessary revenue to finance the tax cut.

Table 3

Economic Effects of Shifting from a Hypothetical 25% Proportional Income Tax to a 22.5% Labor Income Tax, 25% Capital Income Tax, and $900 Lump-Sum Tax

Income Class	Percent of Families	Static		Dynamic		
		Percent Change in After-Tax Income		Percent Change in Pre-Tax Income	Percent Change in After-Tax Income	Percent Change in Utility (Consumption Equivalent)[1]
		Without Financing	With Financing			
Bottom Quintile	20.0	2.9	-12.3	1.0	-11.4	-22.4
Second Quintile	20.0	2.7	-3.2	0.9	-2.3	-2.9
Middle Quintile	20.0	2.5	-0.9	0.8	-0.1	-0.6
Fourth Quintile	20.0	2.4	0.3	0.8	1.1	0.7
Top Quintile	20.0	2.3	1.4	0.9	2.3	1.8
All	**100.0**	**2.3**	**0.3**	**0.9**	**1.1**	**-4.7**

Source: CEA calculations.

Note: Aggregate economic impacts are computed using the macroeconomic model of Mankiw and Weinzierl (2004, 2006). Values for individual families are assumed to change by the same percentage as the aggregate values. The distribution of income is derived from the 2010 IRS Statistics of Income Public Use File. See text for additional details.

[1] At the family level, the consumption equivalent utility increase is the percent change in consumption (assuming labor supply remains unchanged at the baseline level) that would yield the utility level realized in the alternative policy scenario. Percent change for an income class is the simple average of the percent change across families. Utility is computed as log(after-tax income) - n^(1+1/σ), where n is the value of labor supply generated by the Mankiw-Weinzierl model (assuming an isoelastic specification of labor disutility).

Another way to summarize the impacts of the policy is shown in Table 4. Overall, nearly all families would see their pre-tax incomes go up and all working families would see their leisure time fall, while two-thirds of families would see their taxes go up.[7] As a result, 52 percent of families would see their after-tax income go down and nearly 60 percent of families would see their utility go down. These results do not provide a specific judgment on the merits of the policy, which would depend on one's ethical preferences. If you believe that a hypothetical proportional tax with no exemption embodies too much redistribution, then the shift to more lump-sum taxation would be a positive. (Note, that proposals for a flat tax are considerably more progressive than the starting point in this example in that they have a large exemption and thus rising average tax rates.) The minimal suggestion I am making is that policymakers should see analysis like Table 3, allowing them to apply their own social welfare functions to evaluate it.

[7] Families that have only transfer income (and no capital or labor income) will experience no change in pre-tax income. Families that have no labor income under the baseline proportional income tax are likewise assumed to have no labor income under the alternative policy scenario and thus leisure is unchanged. All families with labor income are assumed to increase hours worked (and thus reduce leisure).

Table 4

Economic Effects of Shifting from a Hypothetical 25% Proportional Income Tax to a 22.5% Labor Income Tax, 25% Capital Income Tax, and $900 Lump-Sum Tax

Income Class	Static		Dynamic									
	After-Tax Income		Taxes		Pre-Tax Income		Leisure		After-Tax Income		Utility[1]	
	Percent with Increase	Percent with Decrease	Percent with Increase	Percent with Decrease	Percent with Increase	Percent with Decrease	Percent with Increase	Percent with Decrease	Percent with Increase	Percent with Decrease	Percent with Increase	Percent with Decrease
Bottom Quintile	0.0	100.0	100.0	0.0	95.7	0.0	0.0	90.8	0.0	100.0	0.0	100.0
Second Quintile	0.0	100.0	100.0	0.0	94.8	0.0	0.0	87.2	0.0	100.0	0.0	100.0
Middle Quintile	26.5	73.5	87.4	12.6	95.7	0.0	0.0	84.0	67.6	32.4	48.0	52.0
Fourth Quintile	70.7	29.3	31.9	68.1	97.1	0.0	0.0	84.9	75.1	24.9	71.9	28.1
Top Quintile	84.4	15.6	17.4	82.6	99.0	0.0	0.0	91.3	88.1	11.9	84.2	15.8
All	36.3	63.7	67.3	32.7	96.4	0.0	0.0	87.7	46.2	53.8	40.8	59.2

Source: CEA calculations.

Note: Aggregate economic impacts are computed using the macroeconomic model of Mankiw and Weinzierl (2004, 2006). Values for individual families are assumed to change by the same percentage as the aggregate values. The distribution of income is derived from the 2010 IRS Statistics of Income Public Use File. See text for additional details.

[1] Utility is computed as log(after-tax income) - n^(1+1/σ), where n is the value of labor supply generated by the Mankiw-Weinzierl model (assuming an isoelastic specification of labor disutility).

A more ambitious suggestion would be to present summary statistics for the results, which Table 5 does in the form of the impact on different measures of income and aggregate welfare. Utility is computed as $\log(c) - n^{1+1/\sigma}$, a choice motivated by the similar utility function underlying the Mankiw-Weinzierl analysis. The final two welfare measures presented in the table incorporate a social welfare function that indicates a stronger preference for equity than the pure utilitarian metric, which simply sums utility values for each individual. Percent changes in these welfare measures have little intuitive meaning and the deltas between policy scenarios are not comparable across measures. However, each measure can be used to rank different policies. In this case, both measures indicate that the *ex-ante* policy of a 25 percent proportional tax with no exemption was preferable to the *ex-post* 22.5 percent proportional tax combined with a lump sum tax of $900 per household.

Table 5

Economic Effects of Shifting from a Hypothetical 25% Proportional Income Tax to a 22.5% Labor Income Tax, 25% Capital Income Tax, and $900 Lump-Sum Tax

	Baseline (25% Flat Tax)	Alternative (22.5% Labor Tax + $900 Lump Sum)	**Percent Change**
Income			
Mean Pre-Tax Income	63,122	63,690	**0.9**
Mean After-Tax Income	50,221	50,788	**1.1**
Log After-Tax Income	10.2	10.1	**-1.0**
Welfare			
Mean Consumption Equivalent Utility Increase[1]			**-4.7**
Mean Utility	10.00	9.89	
Mean Log of (Utility + 1)	2.39	2.36	
Mean CRRA Transformation of (Utility + 1) with $\theta = 2$[2]	0.91	0.90	

Source: CEA calculations.

Note: Aggregate economic impacts are computed using the macroeconomic model of Mankiw and Weinzierl (2004, 2006). Values for individual families are assumed to change by the same percentage as the aggregate values. The distribution of income is derived from the 2010 IRS Statistics of Income Public Use File. Utility is computed as log(after-tax income) - n^(1+1/σ), where n is the value of labor supply generated by the Mankiw-Weinzierl model (assuming an isoelastic specification of labor disutility). See text for additional details.

[1] At the family level, the consumption equivalent utility increase is the percent change in consumption (assuming labor supply remains unchanged at the baseline level) that would yield the utility level realized in the alternative policy scenario.

[2] CRRA refers to the constant relative risk aversion functional form: (u^(1-θ)-1)/(1-θ).

Generalizing the Insights From the Simple Illustration

The potential for output-increasing tax reforms to reduce welfare is not unique to the highly stylized case considered here: a reduction in a proportional labor taxes financed with a lump-sum tax. Many revenue-neutral tax reforms have these features, since it can be hard to make large reductions in distortions affecting work-leisure choices without making the tax code much less progressive. Some reforms purport to lower tax rates on work through the canonical distribution-neutral "broader base and lower rate" tax reforms. But in general, such reforms reduce the effective tax rate on work by much less than the headline numbers would suggest. For example, speaking very loosely, eliminating the State and local tax deduction and using the savings to cut income tax rates would mean that a filer currently claiming the deduction would face about a 6 percent lower tax rate on all of their income, rather than a slightly higher tax rate on 93 percent of their income. In other words, the return to an additional dollar worked would be little-changed. Curbing the charitable deduction and using the savings to reduce the tax rate would have a similar effect—a person could keep more of their additional income from work but the reason people work is to consume and the price of their consumption (i.e., the after-tax cost of charitable contributions) would go up. The exact magnitude of the offset depends on the income

elasticity of the item underlying any particular tax expenditure and the heterogeneity across filers in the use of the tax expenditure, but the general point holds—and, in fact, the shape of a distributional table is not a bad guide to the effective marginal tax rate people face. (As discussed above, removing tax expenditures could have welfare benefits in terms of improved consumption without increasing growth through higher labor supply.)

Moreover, the importance of looking beyond output measures also applies, perhaps even more strongly, to revenue-*losing* tax cuts. This is because these tax cuts will eventually have to be financed, and such financing needs to be considered in any welfare analysis because it represents a future cost to *someone*.

Of course, not all tax reforms have this character. A business tax reform could be distribution-neutral and growth-enhancing. A tax reform that involves windfall gains from old capital or work subsidies for low-income workers could be both progressive and growth-enhancing. I am not arguing that tax changes can only improve growth at the expense of equity, just that we should be showing policymakers what—if any—tradeoffs actually exist in practice. Moreover, to the degree that dynamic analysis is conducted in forward-looking models with a particular financing assumption (e.g., lump-sum financing)—a framework the official scorekeepers appear to be moving away from but that is still present in many more academic models—the results should not be presented alongside distributional analysis that implicitly or explicitly uses another financing assumption (e.g., no financing) to give policymakers the misleading impression that they can get the growth effects without the distributional impact.

Implications for Tax Policy in the Real World

So far, I have largely relied on examples to make a point that, I believe, is indisputably true: that we should be evaluating policy proposals based on a rich assessment of their welfare impacts and not based on a narrow focus on output. I have also tried to convince you that these examples are illustrative of more general points: that income and welfare can move in different directions; that certain reforms increasing output often reduce the incomes of low-income families; and that financing is critically important. Crucially, none of these general points depends on any particular model of how the economy operates or any particular social welfare function.

I want to end with an important implication of the preceding analysis for thinking about tax reform in the real world, which is that the distributional effects of tax changes can easily be an order of magnitude higher than their effects on growth. This is not just an empirical observation, but emerges from the underlying economics of the models used for dynamic analysis. As discussed earlier, when a worker increases her labor supply in response to a reduction in the tax rate, she both earns more and thus benefits from higher consumption *and* works more and thus suffers due to lost leisure. For an incremental tax reform, these two factors exactly offset, and for a larger tax reform the impact of the behavioral change on the worker's utility will remain small relative to the impacts of the inframarginal tax changes and effects on the government budget deficit. Thus, the welfare effects of a change in tax policy for any family depend on two inputs: the static change in tax liability resulting from the reform and—in the case of a reform that does not achieve budget balance—the allocation of the eventual financing. At the time a proposal is

evaluated by analysts, the allocation of any required financing is inherently speculative, not only because the allocation rule itself is unknown but also because the amount of financing required will generally depend on the form the financing takes. Thus, the welfare-relevant information an analyst can provide is exactly that captured in a traditional static distribution table. Importantly, this observation does not depend on any particular social welfare function. As long as your assessment of the proposal will rely on the impacts that a proposal has on families' utilities in one way or another, the information you need for that assessment will be found in the static distribution table.[8]

The relative magnitude of distribution and growth effects is borne out by the contrast between the experience of the last eight years and the estimated effects of growth-enhancing tax reforms according to a wide array of analyses. First, note that most credible dynamic analyses show that realistic tax reforms have only a modest impact on growth. For example, Treasury's analysis of the 2005 Bush Tax Reform Panel's Simplified Income Tax plan found it would increase national income by 0.2 to 0.9 percent in the long run (Table 6). Similarly, Treasury's analysis of permanent extension of the 2001 and 2003 tax cuts found they would either increase real GNP by less than 1 percent or decrease real GNP by less than 1 percent depending on financing.

Highly stylized tax reforms—reforms that are easily implemented in an economic model but much more difficult to legislate—can achieve larger growth effects. For example, Altig et al. (2001) find that a proportional consumption tax can increase output by 9 percent in the long run. However, this increase is generated by what is effectively a one-time tax on existing wealth at the time of the reform combined with a sharply regressive redistribution of the tax burden. When transition relief is provided for old capital and a standard deduction is provided in computing the tax, the long-run growth effect is reduced to 2 percent. (Moreover, even with this positive growth effect, the majority of families are made worse off by this reform in the long run.)

[8] Issues of simplification are beyond the scope of this discussion. As discussed in Furman (2016), these benefits can often be overstated, since reducing but not repealing tax expenditures does not increase simplicity; much of the complication in the tax code comes from the measurement of income, not deductions from it, so that there is a tradeoff between efficiency and simplicity; and tax preparation for the median household is not particularly time-consuming or costly.

Table 6

Select Estimates of the Effect of Tax Reform on the Level of Output

Source	Policy Change	Short-Run	Long-Run
Gravelle (2014)	Stylized Reform: 20% Reduction in Income Tax Rates	n. r.	0.7 - 4.0
JCT (2014)	Camp Plan	0.1 - 1.6	n. r.
Treasury (2006b)	President's Advisory Panel on Tax Reform		
	Simplified Income Tax	0.0 - 0.4	0.2 - 0.9
	Growth and Investment Tax	0.1 - 1.9	1.4 - 4.8
	Progressive Consumption Tax	0.2 - 2.3	1.9 - 6.0
Treasury (2006a)	Permanent Extension of the 2001/2003 Tax Cuts		
	Financed with Future Spending Cuts	0.5	0.7
	Financed with Future Tax Increases	0.8	(0.9)
JCT (2005)	Cut in Federal Individual Income Tax Rates		
	(4.0% in first decade, 2.9% thereafter)		
	Not financed	0.0 - 0.5	(0.2) - (0.6)
	Financed with Future Spending Cuts	0.1 - 0.3	0.3 - 0.4
	20% Cut in Federal Corporate Tax Rate		
	Not financed	0.2 - 0.4	0.0 - 0.3
	Financed with Future Spending Cuts	0.2 - 0.4	0.5 - 0.9
CBO (2004)	10% Cut in Federal Individual Income Tax Rates		
	Financed with Future Spending Cuts	0.2	(0.4)
	Financed with Future Tax Increases	0.3	(2.1)
Altig et al. (2001)	Stylized Revenue-Neutral Tax Reforms		
	Proportional Consumption Tax	6.3	9.4
	Flat Tax with Transition Relief	0.5	1.9

n.r. = Not reported. Red indicates negative values.

Note: Output measure is (in order of preference if multiple measures are reported) national income, real gross national product, and real gross domestic product. Time period for short-run effects varies across studies, but (in most cases) is an average over several years in the first decade. Long-run effects typically reflect estimates of the change in the steady state level of output.

Now contrast these impacts with the direct effects that policy changes can have on after-tax incomes through the distributional channel. Specifically, consider the impact that the changes in tax policy over the last 30 or so years have had after-tax incomes. Reductions in average tax rates have increased after-tax incomes in the bottom three quintiles by roughly 6 percent (Table 7). These increases compare favorably with the more modest and realistic tax reforms shown in the table above.

Table 7
Change in After-Tax Income due to Changes in Average Tax Rates by Income Percentile, 1986 to 2013

Income Percentile	Net of Tax Rate, 1986	Net of Tax Rate, 2013	**Percent Change**
0-20	90.7	96.7	**6.6**
20-40	85.5	91.6	**7.1**
40-60	81.9	87.2	**6.5**
60-80	79.4	83.0	**4.5**
80-90	77.2	79.3	**2.7**
90-95	76.5	77.0	**0.7**
95-99	76.3	73.7	**-3.4**
99-100	75.3	66.0	**-12.4**

Source: CBO (2016); CEA calculations.

Note: Net of tax rate is 100 minus the average tax rate. Change in after-tax income due to changes in average tax rate is the percent change in the net of tax rate.

Of course, this analysis does not imply that growth does not matter. Growth is a critical aim of economic policy. And there is a wide array of policies that are growth-enhancing and either progressive or neutral with respect to the distribution of income, such as the progressive expansion of the refundable Child Tax Credit mentioned above. The primary purpose of my remarks today has been to draw out the distinctions between growth and welfare, and to emphasize the value of the traditional static distribution table in informing our assessments of the welfare impacts of potential changes in tax policy.

However, I also think our notions of what contributes to growth need to be broadened. For example, beyond the arguments I have been making today, we need to ask how dynamic analysis should incorporate the effects of reducing liquidity constraints for low-income households or increasing the risk tolerance of middle-income households. But those are topics for another day.

References

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters, and Jan Walliser. 2001. "Simulating Fundamental Tax Reform in the United States." *American Economic Review* 91 (3): 574-595.

Donovan, Shaun. 2015. "'Dynamic Scoring' Is Not the Answer." The White House Blog. January 6.

Furman, Jason. 2016. "What Progressive Tax Reform Should Look Like." *Democracy: A Journal of Ideas*. April 4.

Furman, Jason, and Matthew Fiedler. 2016. "The Cadillac Tax – A Crucial Tool for Delivery-System Reform." *New England Journal of Medicine,* 374 (11): 1008-1009.

Gravelle, Jane G. 2014. "Dynamic Scoring for Tax Legislation: A Review of Models." Congressional Research Service.

Hall, Robert E. and Alvin Rabushka. 1995. *The Flat Tax*. Stanford, CA: Hoover Institution Press.

Mankiw, N. Gregory. 2013. "Defending the One Percent." *Journal of Economic Perspectives* 27 (3): 21-34.

Mankiw, N. Gregory and Matthew Weinzierl. 2004. "Dynamic Scoring: A Back-of-the-Envelope Guide." NBER Working Paper No. 11000.

_____. 2006. "Dynamic Scoring: A Back-of-the-Envelope Guide." *Journal of Public Economics* 90 (8): 1415-1433.

U.S. Congress, Joint Committee on Taxation (JCT). 2014. "Macroeconomic Analysis of the 'Tax Reform Act of 2014'." Report.

U.S. Department of the Treasury, Office of Tax Analysis. 2006a. "A Dynamic Analysis of Permanent Extension of the President's Tax Relief." Report.

_____. "A Summary of the Dynamic Analysis of the Tax Reform Options Prepared for the President's Advisory Panel on Federal Tax Reform." Report.