Journal of Public Economics 187 (2020) 104198

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Measuring economic efficiency using inverse-optimum weights☆

Nathaniel Hendren

Harvard University, United States of America

ARTICLE INFO

Article history:
Received 5 December 2019
Received in revised form 12 April 2020
Accepted 20 April 2020
Available online 6 June 2020

ABSTRACT

This paper provides a method to measure the traditional Kaldor-Hicks notion of "economic efficiency" when taxes affect behavior. In contrast to traditional unweighted surplus, measuring efficiency requires weighting individual benefits (or surplus) by the marginal cost to the government of providing a $1 transfer at each income level. These weights correspond to the solution to the "inverse-optimum" program in the optimal tax literature: they are the social planning weights that would rationalize the status quo tax schedule as optimal. I estimate the weights using the universe of US income tax returns from 2012. The results suggest that measuring economic efficiency requires weighting surplus accruing to the poor roughly 1.5–2 times more than surplus accruing to the rich. This is because $1 of surplus to the poor can be turned into roughly $1.5–$2 of surplus to the rich by reducing the progressivity of the tax schedule. Following Kaldor and Hicks' original applications, I compare income distributions over time in the US and across countries. The results suggest US economic growth is 15–20% lower due to increased inequality than is suggested by changes in GDP. Because of its higher inequality, the U.S. is unable to replicate the income distribution of countries like Austria and the Netherlands, despite having higher national income per capita.

1. Introduction

Suppose an alternative environment offers benefits or "surplus" $s(y)$ for each person with income y. Some people may be better off in the alternative environment, $s(y)>0$; others may be worse off, $s(y)<0$. Deciding whether the alternative environment is better than the status quo therefore requires resolving these interpersonal comparisons: how should society weight the gains to the winners against the losses to the losers?

A common method for resolving these tradeoffs posits a set of social welfare weights, $\chi(y)$, for each level of income y (e.g. Saez and Stantcheva (2016)). For example, these weights may be decreasing in income so that they capture a preference for equity. This approach would then ask whether the weighted average of surplus is positive, $E[s(y)\chi(y)]>0$. However, the downside of this approach is that it generates conclusions that depend on the social welfare weights. Because these weights reflect ethical and philosophical tradeoffs about which there is no consensus, this approach can fail to generate universal agreement about whether the alternative environment should be preferred to the status quo.

Eight decades ago, Kaldor (1939) and Hicks (1940) proposed a method to resolve this problem. Instead of appealing to set of social welfare weights, they proposed modifying the environments with transfers. They show that an appropriate set of transfers removes the need for interpersonal comparisons in the first place; instead they rely only on the Pareto principle. Kaldor (1939) noted that when $E[s(y)]>0$ one could construct a modified alternative environment that includes compensating transfers so that the winners compensate the losers. This modified alternative can be preferred to the status quo without requiring a particular set of social welfare weights – everyone would be better off. Analogously, when $E[s(y)]<0$, Hicks (1940) noted that one could construct a modified status quo environment in which transfers make everyone better off relative to the alternative environment – again, one need not appeal to a set of social welfare weights. These two conceptual experiments motivated unweighted surplus, $E[s(y)]$, as a normative measure of economic efficiency.

☆ This paper is a revised version of a paper that previously circulated under the titles, "The Inequality Deflator: Interpersonal Comparisons without a Social Welfare Function" and "Efficient Welfare Weights". I am deeply indebted to conversations with Louis Kaplow for the inspiration behind this paper, and to Sarah Abraham, Alex Bell, Alex Olssen, Peter Ruhm, and Evan Storms for excellent research assistance. I also thank Daron Acemoglu, Raj Chetty, Amy Finkelstein, Ben Lockwood, Henrik Kleven, Patrick Kline, Kory Kroft, Matthew Notowodigdo, Jim Poterba, Emmanuel Saez, Matthew Weinzierl, Glen Weyl, Ivan Werning, and Floris Zoutman, along with seminar participants at Berkeley, Harvard, MIT, Michigan, and Stanford for very helpful comments. The opinions expressed in this paper are those of the author alone and do not necessarily reflect the views of the Internal Revenue Service or the U.S. Treasury Department. This work is a component of a larger project examining the effects of tax expenditures on the budget deficit and economic activity, and this paper in particular provides a general characterization of the welfare impact of changes in tax expenditures relative to changes in tax rates. The empirical results derived from tax data that are reported in this paper are drawn from the SOI Working Paper "The Economic Impacts of Tax Expenditures: Evidence from Spatial Variation across the U.S.", approved under IRS contract TIRNO-12-P-00374. I gratefully acknowledge funding from the National Science Foundation (CAREER1653686).

E-mail address: nhendren@fas.harvard.edu.

https://doi.org/10.1016/j.jpubeco.2020.104198

Despite its reliance on the Pareto principle, the Kaldor-Hicks test for efficiency is mathematically equivalent to the social welfare weight approach with $\chi(y) = 1$ for all y. For this reason, the Kaldor-Hicks definition of efficiency is often criticized for its lack of consideration for distributional equity. The reliance on the Pareto principle arguably only applies if the transfers are carried out. In practice governments cannot generally implement the type of non-distortionary individual-specific lump-sum transfers envisioned in the experiments above. Taxes are imposed on observable choices like incomes that respond to taxes and transfers (Mirrlees (1971)). In fact, Kaldor and Hicks themselves argued that correctly measuring economic efficiency requires accounting for these costs:

> "*Since almost every conceivable kind of compensation (re-arrangement of taxation, for example) must itself be expected to have some influence on production, the task of the welfare economist is not completed until he has envisaged the total effects...If, as will often happen, the best methods of compensation feasible involve some loss in productive efficiency, this loss will have to be taken into account*" (Hicks (1939), p712).

This paper provides a method to measure Kaldor-Hicks efficiency that accounts for the distortionary cost of redistribution across different income levels. To do so, I show that to first order one must weight surplus at each income level \$$y$ by the marginal cost of providing \$1 of welfare to individuals earning near \$$y$, $g(y)$. These weights $g(y)$ are known in existing literature as the "inverse optimum" social welfare weights (see, e.g., Christiansen (1977); Christiansen and Jansen (1978); Blundell et al. (2009); Bargain et al. (2011); Bourguignon and Spadaro (2012); Lockwood and Weinzierl (2016); Zoutman et al. (2013); Bargain et al. (2014); Jacobs et al. (2017)). In other words, a social planner that has planning weights $g(y)$ would find that the status quo tax schedule maximizes its objective function. In addition to revealing these implicit planning preferences, this paper shows that these weighted surplus, $E[s(y)g(y)]$, measures economic efficiency in a way that accounts for the distortionary cost of taxation. When $E[s(y)g(y)] > 0$, there exists an alternative environment with a modified tax schedule that provides a Pareto improvement relative to the status quo. When $E[s(y)g(y)] < 0$, the tax schedule in the status quo can be modified to provide a Pareto improvement relative to the alternative environment. In this sense, measuring economic efficiency as $E[s(y)g(y)]$ implements the conceptual experiments of Kaldor and Hicks in a manner that accounts for the distortionary cost of taxation.

What do the inverse-optimum weights, $g(y)$, look like in the US? If taxes did not affect behavior, the weights would be 1 at all levels of income: the cost of providing a \$1 tax cut to those earning near \$$y$ would be simply \$1. But, this cost differs from \$1 when individual incomes respond to the tax cut. For example, those earning below (above) \$$y$ might increase (decrease) their incomes to obtain the additional income. By the envelope theorem, these behavioral responses do not generate a first-order impact on utility, but they do generate a first-order impact on the cost of the tax cut to the government. If taxes are positive (negative), increases in incomes create positive (negative) fiscal externalities that reduce (increase) its cost.

To quantify these responses, I leverage the derivation provided in Jacobs et al. (2017) that shows the impact of the behavioral response to taxation on the government budget can be expressed as a function of (a) the joint distribution of taxable income and marginal tax rates and (b) a set of behavioral elasticities governing the response of income to changes in taxation.[1] I use the universe of US income tax returns from 2012 to estimate this joint distribution, and I begin by providing bounds on the weights (without assuming a magnitude of the behavioral response to taxation). I show that the shape of the income distribution - in particular the local Pareto parameter of the income distribution - plays a key role in determining the extent to which the weights are above or below 1.[2] The Pareto parameter rises from near −1 at the bottom of the income distribution to near 2 at the top of the income distribution, crossing zero around the 60th quantile of the income distribution (around \$43 K in ordinary income). This means that weights are generally above one for those with incomes below \$43 K, and below one for those with incomes above \$43 K. Regardless of the size of the behavioral response to taxation, it is more costly to provide \$1 to the poor than to the rich. Thus, these bounds suggest it is efficient to weight surplus to the poor more than to the rich.[3] Intuitively, it is more costly to move an additional \$1 from the top to the bottom of the income distribution through additional redistribution than it is to move \$1 from the bottom to the top of the income distribution through reduced redistribution.

Next, I construct point estimates using existing estimates of taxable income elasticities. The baseline specification suggests a \$1 tax cut to those with high incomes from a reduction in marginal tax rates costs around \$0.65. At the other end of the income distribution, the estimates suggest that expansions of the earned income tax credit (EITC) by \$1 to low earners has a fiscal cost of around \$1.15 because additional transfers cause individuals to adjust their earnings to maximize their tax credits. This means that the the weights decline from around 1.15 at the bottom of the income distribution to around 0.65 at the top. In other words, \$1 to the poor can be turned into roughly \$1.77 to the rich through modifications to the tax schedule.[4] As a result, it is efficient to weight surplus to the poor roughly twice as much as surplus to the rich. And, social welfare weights that place roughly 1.77 times more weight on the poor relative to the rich would rationalize the status quo tax schedule as optimal.

Motivated by the original applications in the work of Kaldor (1939) and Hicks (1940), I apply the weights to two sets of comparisons of income distributions.[5] First, I construct distributionally-adjusted measures of economic growth in the US. As is widely documented, growth in the US has been unequal across the income distribution. Because it is costly to redistribute from rich to poor, distributionally-adjusted measures of economic growth are 15–20% lower. If economic growth were redistributed equally across the income distribution, US per capita growth would go down by 15–20%. Extrapolating across all economic growth between 1979 and 2012 suggests an increase in distributionally-adjusted growth of \$15K, in contrast to aggregate growth of \$18K. Multiplying by 119M households in the US suggests a social cost of increased income inequality in the US since 1979 of roughly \$400B.

Second, I compare the distribution of incomes across countries. Broadly, orderings of countries by unweighted mean incomes tend to yield the same conclusions as ordering by weighted incomes. But, there are several exceptions. Most notably, the income distributions of Austria and New Zealand would be preferred relative to the US income distribution, despite having a lower per capita income. Although the US has higher mean income, it is unable to replicate the distribution of income offered in those countries through modifications in the tax schedule.

[1] I extend the result in Jacobs et al. (2017) to allow for the presence of multiple tax schedule for those at the same taxable income level, as occurs in the US.

[2] The Pareto parameter is given by $-(1+\frac{yf'(y)}{f(y)})$ where $f(y)$ is the density of the income distribution.

[3] This is also consistent with reduced form evidence that suggests it costs the government more to provide an additional dollar to the poor than to the rich. For example, reductions in top income taxes tend to increase incomes, which reduces the net cost to the government (Saez et al. (2012). In contrast, additional transfers to the bottom of the income distribution can increases the cost to the government (Hendren (2016), Hotz and Scholz (2003) and Chetty et al. (2013)).

[4] These weights estimated from tax data are consistent with results in Lockwood and Weinzierl (2016) who estimate the solution to the inverse optimum program in the U.S. using aggregated data from the Congressional Budget Office.

[5] In Appendix G, I also discuss the implications for the welfare impacts of economic policies that target particular regions of the income distribution.

This is not the first paper to recognize that incorporating the distortionary cost of transfers leads to a modification of the Kaldor and Hicks compensation principle. Early discussions of these ideas are found in Christiansen (1981), and later by Kaplow (2004) and others, who discuss winners compensating losers through modifications to the tax schedule. The theoretical analysis in this paper is most closely related to Coate (2000), who proposes an approach that incorporates the costs of redistribution into the Hicks criterion by comparing the policy to feasible alternatives such as distortionary redistribution through the tax schedule. Coate (2000) writes: *"An interesting problem for further research would be to investigate whether the efficiency approach might be approximately decentralised via a system of shadow prices which convey the cost of redistributing between different types of citizens."* This paper shows that the weights corresponding to the solution to the inverse optimum program in the optimal tax literature provide these appropriate shadow prices. In other words, the inverse-optimum weights provide a first-order method for measuring economic efficiency as originally envisioned by Kaldor (1939) and Hicks (1940).

The rest of this paper proceeds as follows. Section 2 provides the theoretical setup including the status quo and alternative environment, and provides a general definition of the marginal cost of taxation (inverse optimum weights). Section 3 illustrates how the weights implement the modified Kaldor-Hicks efficiency experiments. Section 4 uses the derivation in Jacobs et al. (2017) to represent these weights using the distribution of income, tax rates, and behavioral elasticities. Section 5 provides estimates of the joint distribution of income and tax rates and discusses bounds on the shape of the weights. Section 6 provides point estimates by calibrating behavioral elasticities. Section 7 applies the weights to the comparison of income distributions over time in the US and across countries. Section 8 discusses limitations of the approach, and Section 9 concludes.

2. Model

This section develops a model that is used to define two key variables that will be important for implementing the Kaldor-Hicks tests for efficiency. First, for a given alternative environment, I use the model to define each person's willingness to pay for this environment. Second, I define the marginal cost to the government of providing a $1 transfer to those earning a given income level, y. As noted in the introduction, these weights are also known in existing literature as the solution to the inverse optimum program.

2.1. Willingness to pay

I consider an economy with a unit mass of agents, indexed by θ. There is a status quo environment and an alternative environment. The alternative environment could be a world with greater spending on a public good, a more progressive tax schedule, or the distribution of income offered by another country. This latter case of comparisons of income distributions was the motivating comparison considered in Kaldor (1939) and Hicks (1940), which I return to below in Section 7.

In the status quo environment, agents consumption, c, and earnings, y. I index agents by their type θ and allow each θ to have a different utility function, $u(c,y;\theta)$, over consumption and earnings. I do not impose restrictions on the distribution of θ. Agents choose c and y to maximize utility subject to a budget constraint,

$$c \le y - T(y) + m$$

where $T(y)$ is the taxes paid on earnings y and m is additional income beyond earnings.[6] With a slight abuse of notation, I let $c(\theta;T(\circ))$ and $y(\theta;T(\circ))$ denote the resulting choices of type θ in the status quo environment with tax schedule $T(\circ)$.[7]

Let $v^0(\theta;T(\circ))$ denote the utility level obtained by type θ in the status quo environment when facing tax schedule $T(\circ)$. And, given a utility level v, define the expenditure function $e(v;\theta)$ to be the smallest value of m that is required for a type θ to obtain utility level v in the status quo environment.[8]

Let $u^a(c,y;\theta)$ denote the utility function for type θ in the alternative environment. I do not restrict any feature of the alternative environment – it could contain different wage distributions, better schools, less traffic, better restaurants, or simply different scenery – any of which can affect the level of u^a for any individual θ. I also do not restrict that the tax schedule in the alternative environment be the same as the status quo. To that aim, let $T^a(\circ)$ denote the tax schedule in the alternative environment so that the budget constraint in the alternative environment is given by $c \le y - T^a(y) + m$. Define $v^a(\theta;T^a(\circ))$ to be the level of utility obtained and $e^a(v;\theta)$ is the smallest value of m that is required for a type θ to obtain utility level v in the alternative environment. Given the tax schedules in the status quo and alternative environment, individual θ's willingness to pay (equivalent variation) for the alternative environment relative to the status quo is then given by

$$s(\theta) = e\left(v^a\left(\theta;T^a(\circ)\right);\theta\right) - e\left(v^0\left(\theta;T(\circ)\right);\theta\right) \quad (1)$$

where $e(v^0(\theta;T(\circ));\theta) = m$. The value $s(\theta)$ is the amount of additional money a type θ would need in the status quo to be as well off as in the alternative environment.[9]

I make some simplifying assumptions on this surplus function, $s(\theta)$, that are relaxed in Appendix D. In particular, I assume that it does not vary with θ conditional on income, $y(\theta;T(\circ))$. With an abuse of notation, I let $s(y)$ denote the willingness to pay for the alternative environment by a type θ who chooses income $y(\theta)$ in the status quo. And, for simplicity I assume that $s(y)$ is continuous in income, y.

Given $s(y)$, the goal is to answer two questions: (1) can the surplus, $s(y)$, can be replicated through modifications in the tax schedule in the status quo environment (i.e. the experiment in Hicks (1940))? And, (2) does there exists a modification to the tax schedule in the alternative environment that makes everyone better off relative to the status quo (i.e. the experiment in Kaldor (1939))? The answer to these questions will depend on how changes to the tax schedule affect government revenue.

2.2. Marginal cost of taxation (a.k.a. inverse-optimum weights)

This subsection defines the marginal cost of providing a transfer to those with incomes at a given level in the status quo environment. To make this formal, note that government revenue is given by $E[T(y(\theta;T(\circ)))]$ (recall there is a unit mass of individuals). This equals the average amount of taxes collected across the population who choose incomes $y(\theta;T(\circ))$. I assume that $E[T(y(\theta;T(\circ)))]$ is continuously differentiable in $T(\circ)$.[10]

[6] For simplicity, I assume $T(y)$ is the same for everyone. In the empirical implementation, I allow T to vary with individual characteristics, such as the number of dependents, and marital status. See Section E.1.

[7] These choices also depend on m, but I suppress this notation for brevity.

[8] Formally, $e(v;\theta) = \inf\{m | \sup_{c,y}\{u(c,y;\theta) | c \le y - T(y) + m\} \ge v\}$. The standard duality result implies that $e(v^0(\theta;T(\circ));\theta) = m$.

[9] In addition to this equivalent variation definition of willingness to pay, one could also construct a compensating variation measure using the expenditure function in the alternative environment, $cv(\theta) = e^a(v(\theta;T(\circ));\theta) - e^a(v^a(\theta;T^a(\circ));\theta)$. Because the distinction between equivalent and compensating variation is second order (e.g. see Schlee (2013) for a recent discussion of the first-order equivalence of five common conceptualizations of willingness to pay, including compensating and equivalent variation.) and the approach below considers first-order adjustments, it will not be necessary to distinguish between equivalent or compensating variation in the analysis that follows.

[10] Formally I assume that for any function $h(y)$ of taxable income y, let $\tilde{T}_\varepsilon(y) = T(y) + \varepsilon h(y)$. Then $E[\tilde{T}_\varepsilon(y(\theta;\tilde{T}_\varepsilon(\circ)))]$ is continuously differentiable in ε for any function $h(y)$. Note this allows for individual behavioral responses to be discontinuous (e.g. extensive margin responses) – only population-average tax revenue is required to be continuously differentiable.

Now, consider a tax schedule modified tax schedule, $\hat{T}$, that provides \$$\eta$ of a tax cut to those with incomes in a region of width ε near y^*:

$$\hat{T}(y;y^*,\varepsilon,\eta)=\begin{cases}T(y) & if \;\; y\notin\left(y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}\right)\\ T(y)-\eta & if \;\; y\in\left(y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}\right)\end{cases}$$

$\hat{T}$ provides η additional resources to an ε-region of individuals earning between $y^*-\varepsilon/2$ and $y^*+\varepsilon/2$. This is depicted in Fig. 1 for a tax cut of size $\eta=1$.

For a small increase in transfers starting at $\eta=0$, the envelope theorem implies that individuals with earnings between $y^*-\varepsilon/2$ and $y^*+\varepsilon/2$ will be willing to pay η to have a tax schedule given by $\hat{T}(y;y^*,\varepsilon,\eta)$ instead of $T(y)$.[11] However, the marginal cost to the government of this policy per mechanical beneficiary is not equal to \$$\eta$. This is because there is an additional cost (or benefit) to the government due to the behavioral responses to taxation.

To capture this formally, consider the derivative of government revenue with respect to the size of the tax cut, η, evaluated at $\eta=0$, divided by the fraction of mechanical beneficiaries whose income in the status quo is in the ε-region near y^*. Using the notation above, this is given by $\frac{d[E[\hat{T}(y(\theta;\hat{T}(\circ;y^*,\varepsilon,\eta));y^*,\varepsilon,\eta)]]}{d\eta}|\eta=0/\ \Pr\{y(\theta;T(\circ)\in[y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}])\}$. The first term is the marginal cost to the government of providing the tax cut to those earning within the $\varepsilon-$region of y^*. This equals the derivative of the average tax collected from those of each type θ when facing this modified tax schedule, $y(\theta;\hat{T}(\circ;y^*,\varepsilon,\eta))$. This marginal cost is then divided by the size of mechanical beneficiaries whose incomes in the status quo are in the $\varepsilon-$region of y^*, $\Pr\{y(\theta;T(\circ))\in[y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}]\}$. Taking the limit as $\varepsilon\to0$ yields the marginal cost to the government of providing an additional dollar of resources to an individual earning y^*:

$$g(y^*)\equiv\lim_{\varepsilon\to0}\frac{\frac{d\left[E\left[\hat{T}\left(y\left(\theta;\hat{T}(\circ;y^*,\varepsilon,\eta)\right);y^*,\varepsilon,\eta\right)\right]\right]}{d\eta}}{\Pr\left\{y(\theta;T(\circ))\in\left[y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}\right]\right\}}|\eta=0 \qquad (2)$$

To provide intuition for the formula in Eq. (2), suppose individuals did not change their incomes in response to the tax cut so that $y(\theta;\hat{T}(\circ;y^*,\varepsilon,\eta))=y(\theta;T(\circ))$. In this case, the marginal cost to the government of the tax cut is equal to the number of people whose incomes are eligible for the transfer, $\frac{d[E[\hat{T}(y(\theta;\hat{T}(\circ;y^*,\varepsilon,\eta));y^*,\varepsilon,\eta)]]}{d\eta}|\eta=0=\ \Pr\{y(\theta;T(\circ)\in[y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}])\}$ for any ε. When incomes do not respond to changes in taxes, $g(y)=1$. In this sense, the marginal cost $g(y)$ is conceptually the sum of two components, $g(y)=1+FE(y)$, where $FE(y)$ is the "fiscal externality" from the tax cut: it is the impact of behavioral responses to the tax cut to those with incomes near y on government revenue. Note this fiscal externality could be the result of individuals with incomes near y choosing to increase or decrease their incomes, or it could be the result of those with incomes in the status quo very far away from y choosing to "jump" to having an income near y in order to obtain the tax cut. In Section 4 I provide further assumptions that are employed in Jacobs et al. (2017) that enable the fiscal externality to be represented using empirical elasticities.

[11] This is true as long as the incidence of the tax cut falls entirely on the beneficiaries and does not result in changes in wages. For example, if firms respond to the tax cut of \$1 by lowering wages by \$0.50, then the individual would only be willing to pay \$0.50 for a \$1 tax cut. Here, I assume no general equilibrium responses to taxation. Tsyvinski and Werquin (2018) provide a generalization of this approach to allow for general equilibrium responses to taxation.

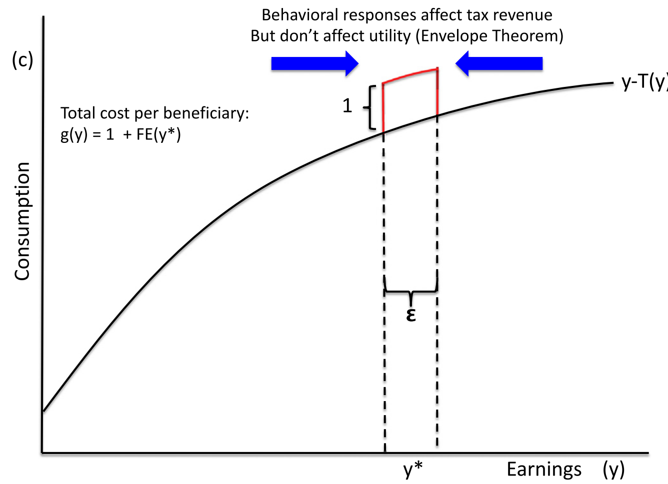

Fig. 1. Tax cut to those earning near y^*. *Notes*: This figure illustrates the modification to the tax schedule that provides a tax cut of \$1 to those with earnings in a region of y^* of width ε. To first order, those whose earnings would lie in $[y^*-\frac{\varepsilon}{2},y^*+\frac{\varepsilon}{2}]$ will value the tax cut at \$1. But, the costs will result from both this mechanical cost and the impact of behavioral responses to the tax cut (loosely illustrated by the blue arrows). So, the total cost per unit of mechanical beneficiary will be $g(y)=1+FE(y)$, where $FE(y)$ is the impact of behavioral responses to the tax cut on government revenue.

Inverse-optimum program

The weights $g(y)$ correspond to the solution to the inverse-optimum program: they are the social welfare weights that rationalize indifference to modifications to the status quo tax schedule.[12] To see this, let $\chi(y)$ denote the impact on social welfare of providing \$1 to an individual earning y. These values of $\chi(y)$ are known as social marginal utilities of income or "generalized social welfare weights" in Saez and Stantcheva (2016) and in general can provide a flexible way of representing a social welfare function.

To see the relationship with the "inverse optimum" weights, let $\chi^*(y)$ denote the particular set of social marginal utilities of income that rationalize the tax schedule as optimal. Next, suppose one provides a small tax cut of \$1 to those earning near y. Those with incomes near y will be willing to pay \$1 for this tax cut, and it will generate a social welfare impact of $1*\chi^*(y)$. However, as outlined in the previous subsection, this transfer will have a cost of $g(y)$. Hence, every dollar of net government spending towards those earning near y will deliver $\frac{\chi^*(y)}{g(y)}$ units of social welfare. If the tax schedule is set to maximize social welfare, this means that the social welfare impact of a tax cut to those earning y' must equal the social welfare impact of a tax cut to those earning y. This means that $\frac{\chi^*(y)}{g(y)}$ must be constant for all y;

$$\frac{\chi^*(y)}{g(y)}=\kappa \quad \forall y$$

Since social welfare weights are only defined up to a constant, $g(y)$ is the unique set of social welfare weights that rationalize the tax schedule as optimal. In this sense, $g(y)$ are the inverse-optimum welfare weights. The next section shows how these weights can be used to measure Kaldor-Hicks efficiency.

[12] See, e. g., Christiansen (1977); Christiansen and Jansen (1978); Blundell et al. (2009); Bargain et al. (2011); Bourguignon and Spadaro (2012); Lockwood and Weinzierl (2016); Zoutman et al. (2013); Bargain et al. (2014); Jacobs et al. (2017) for earlier work in the inverse optimal tax approach.

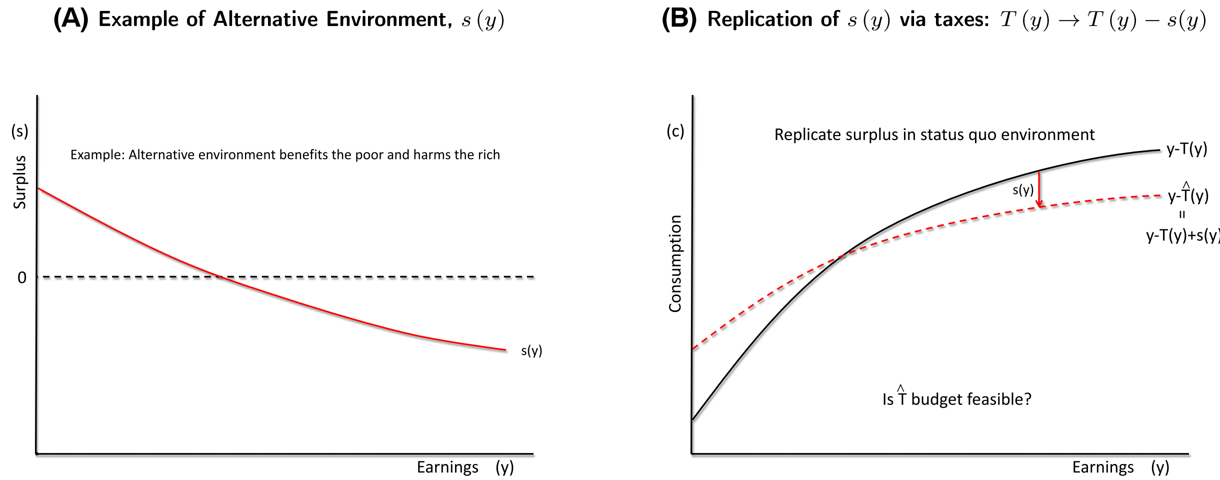

Fig. 2. Hicks (1940) efficiency experiment. *Notes*: This figure illustrates the efficiency experiment of Hicks (1940) for a hypothetical alternative environment. Panel A presents the hypothetical willingness to pay for each person at different points of the income distribution. In this example, those with low incomes prefer the alternative environment, but those with higher incomes prefer the status quo. Panel *B* illustrates modifying the tax schedule in the status quo world to attempt to replicate the surplus offered by the alternative environment. To first order, everyone is indifferent between the alternative environment and the modified status quo with tax schedule $T(y) - s(y)$.

3. Using $g(y)$ to measure economic efficiency

Traditional approaches to measuring Kaldor-Hicks efficiency construct unweighted surplus. $E[s(y)]$. But, this does not incorporate the distortionary cost of the taxes and transfers required to implement the Kaldor and Hicks' experiments. This section shows how one can account to first order for the distortionary cost of redistribution by weighting surplus by $g(y)$ when measuring economic efficiency:

$$S = E[s(y)g(y)] \tag{3}$$

Section 3.1 consider the test of efficiency in the spirit of Hicks (1940): $S > 0$ if and only if one cannot replicate the surplus allocation offered by the alternative environment using modifications to the tax schedule. Section 3.2, implements a first order test of efficiency in the spirit of Kaldor (1939): $S > 0$ if and only if one can modify the tax schedule in the alternative environment to generate a Pareto superior allocation that is preferred by everyone relative to the status quo. Combined, these tests motivate weighting surplus by $g(y)$ to measure economic efficiency.

3.1. Testing for efficiency as defined in Hicks (1940) and Coate (2000)

Can the benefits offered by the alternative environment, $s(y)$, be more efficiently provided through modifications in the tax schedule? To assess this, imagine replacing the current tax schedule, $T(y)$, with a new tax schedule, $\hat{T}(y) = T(y) - s(y)$, that offers a tax cut of size $s(y)$ to those earning y. Fig. 2 provides an illustration. Panel A presents a hypothetical alternative environment that is preferred by the poor but not by the rich. Panel B then modifies the tax schedule from $T(y)$ to $T(y) - s(y)$. To first order, the envelope theorem implies that the tax cut of $s(y)$ is valued at $s(y)$ by those earning y. Therefore, everyone is approximately indifferent between the alternative environment and the status quo environment with the modified tax schedule, as depicted by the dashed red line in Fig. 2, Panel B. The Hicks test for efficiency asks: Is this tax modification in the status quo world feasible?

To first order, the marginal cost of providing $1 of welfare to those earning y is given by $g(y) = 1 + FE(y)$. Therefore, the cost of this tax cut is given by $E[g(y)s(y)]$. If this quantity is positive, then providing surplus $s(y)$ through the tax schedule would not be feasible. Closing the budget constraint by raising taxes on everyone would lead to the blue line in Fig. 3, Panel A. In this sense, the alternative environment would be efficient relative to what is feasible through modifications to the tax schedule in the status quo. In contrast, if $S < 0$, then it is possible to replicate the alternative environment through modifications to the tax schedule. Redistributing the government surplus to everyone equally leads to the blue line in Fig. 3, Panel B, which is preferred by all relative to the alternative environment. In this sense, the alternative environment is efficient if and only if $S > 0$.

The formal version of these statements are valid up to first order, as they rely on the envelope theorem to ensure indifference between the modified status quo (the dashed red line in Fig. 2, Panel B) and the alternative environment. Proposition 1 provides one method of formalizing this idea by considering a scaled surplus function.[13]

Proposition 1. For any $\varepsilon > 0$ define the scaled surplus by $s_\varepsilon(y) = \varepsilon s(y)$ and $S_\varepsilon = E[s_\varepsilon(y)g(y)] = \varepsilon S$. If $S < 0$, there exists an $\tilde{\varepsilon} > 0$ such that for any $\varepsilon < \tilde{\varepsilon}$ there exists an augmentation to the tax schedule in the status quo environment that generates surplus, $s_\varepsilon^t(y)$, that is uniformly greater than the surplus offered by the alternative environment, $s_\varepsilon^t(y) > s_\varepsilon(y)$ for all y. Conversely, if $S > 0$, no such $\tilde{\varepsilon}$ exists.

Proof. See Appendix A.2.

The core insight of Hicks (1940) is that by modifying the status quo through transfers, one can compare the status quo and alternative environment using the Pareto principle, as opposed to relying on a particular social welfare function to make the comparison. I state this result in the following Corollary.

Corollary 1. For any set of (positive) social welfare weights, $\chi(y)$, the augmented status quo environment delivers greater social welfare than the alternative environment, $E[s_\varepsilon^t(y)\chi(y)] > E[s_\varepsilon(y)\chi(y)]$.

Proof. This follows from the fact that $s_\varepsilon^t(y) > s_\varepsilon(y)$ *for all y and that the weights* $\chi(y) > 0$ $\forall y$. ☐

If $S < 0$, then the alternative environment is Pareto dominated by a modification to the tax schedule. In this sense, alternative environments for which $S < 0$ are not desirable. But, what about policies for which $S > 0$? Should these be pursued?

[13] Proposition 1 formalizes the first order approach by scaling the surplus function. Alternatively, one could formalize the approach by directly modeling a continuum of alternative environments in the utility function. For example, suppose a is a continuous number indexing alternative environments (e.g. level of a public goods, trade policy, etc). Let $a = 0$ corresponds to the status quo and assume one can write individuals' utility functions, $u(c,y,a;\theta)$. In this case, one can define $s(y)$ to be individuals marginal willingness to pay out of their own income for a marginal change in a: $s(y) = \frac{\partial u}{\partial a} / \frac{\partial u}{\partial c}$ evaluated at $a = 0$. In this case, a modification to the tax schedule can make everyone better off relative to a world with a slightly higher value of a if and only if $E[g(y)s(y)] < 0$.

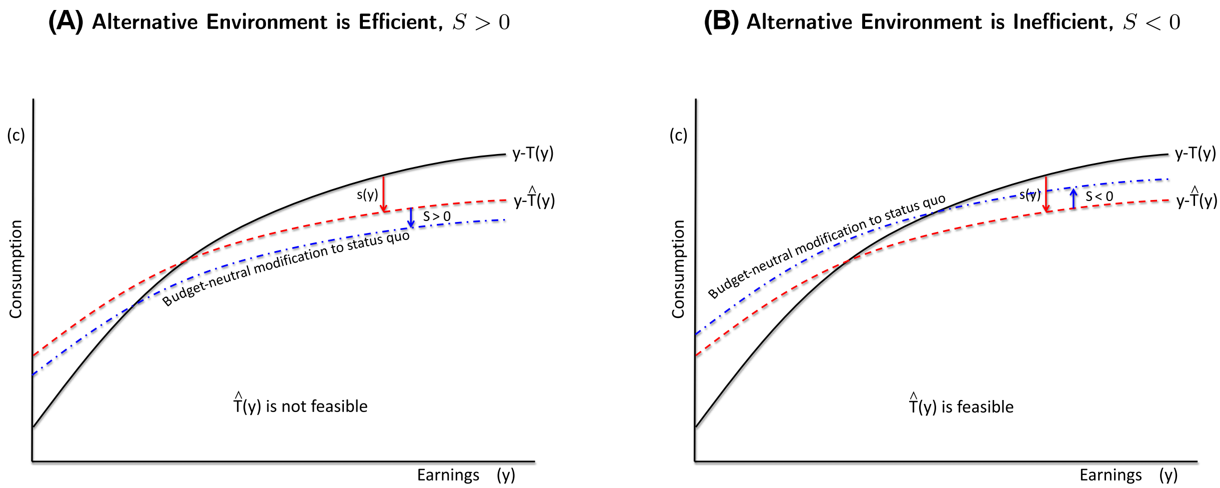

Fig. 3. Testing for Hicks efficiency. *Notes*: This figure illustrates the efficiency test of Hicks (1939). The blue line illustrates the conceptual after-tax income that is feasible through modifications to the tax schedule but has the same distributional incidence as the alternative environment. Panel A illustrates the case in which the modified status quo tax schedule would deliver lower welfare to all points of the income distribution, so that the alternative environment is efficient relative to the status quo, $S > 0$. In contrast, Panel B illustrates the case in which replicating the surplus offered by the alternative environment through the tax schedule leads to higher welfare for all, so that the alternative environment is inefficient.

Armed with only the result in Proposition 1, it is unclear. While Hicks (1940) originally suggested yes, moving to the alternative environment does not generate a Pareto improvement relative to the status quo. Rather, it generates a Pareto improvement relative to a modified status quo that attempts to replicate the distributional incidence of the alternative environment. Actually moving to the alternative environment would generate winners and losers. Hence, $S > 0$ suggests it is a useful policy to consider (it's an "efficient" policy in the sense of Coate (2000)). But, it is not clear whether it is desirable relative to the status quo if $s(y) < 0$ for some y.

In order to provide guidance in the case when efficient surplus is positive, it is useful to consider a different conceptual experiment: that of Kaldor (1939).

3.2. Testing for efficiency as defined in Kaldor (1939)

When can everyone be made better off relative to the status quo environment? Consider modifying the tax schedule in the alternative environment, $T^a(y)$, so that the winners compensate the losers, $T^a(y) \to T^a(y) + s(y)$. Fig. 4 presents this modified tax schedule in the alternative environment. Those with incomes y are better off by $s(y)$ relative to the status quo. The dashed red line in Fig. 4 taxes back these gains. The envelope theorem suggests that to first order individuals earning y in the alternative environment are worse off by $s(y)$ when we tax back these benefits. Everyone is approximately indifferent between the status quo environment and the alternative environment with the modified income tax schedule. Therefore, the question becomes: Is this modification to the tax schedule in the alternative environment budget feasible?

To first order, the modification to the tax schedule generates revenue $S^a = E[g^a(y)s(y)]$, where $g^a(y)$ is the cost to the government of providing \$1 to those earning near \$$y$ in the alternative environment. If $S^a > 0$, a modified alternative environment in which the winners compensate the losers through modifications to the tax schedule can make everyone better off relative to the status quo.

To make these "first order" statements precise, it is also helpful to again consider an $\varepsilon-$scaled alternative environment that delivers $s_\varepsilon(y(\theta)) = \varepsilon s(y(\theta))$ of surplus to each type θ. In this hypothetical alternative environment, I let $g_a^\varepsilon(y^a(\theta))$ denote the marginal cost of taxation. In practice, S_ε^a could differ from S because the marginal cost of a tax cut may differ in the alternative and status quo environment. If this is the case, it could be that the alternative environment dominates all feasible modifications to the status quo tax schedule ($S > 0$) but there does not exist a modified alternative environment that delivers a Pareto improvement relative to the status quo ($S^a < 0$). But, many applications involve sufficiently small changes to the structure of the economy, in which case it seems reasonable to assume that the marginal cost of taxation for each type θ is similar in the status quo and alternative environments, $g^a(y^a(\theta)) \approx g(y(\theta))$. I state this formally in Assumption 1.

Assumption 1. Let $y^a(\theta)$ denote the income of type θ in the alternative environment scaled by ε. The marginal cost of taxation for each type θ is the same in the alternative environment as in the status quo, $g_\varepsilon^a(y^a(\theta)) = g(y)$ for all $\varepsilon \in (0,1]$.

If Assumption 1 holds, then $S > 0$ provides a first-order test of whether those with $s(y) > 0$ can compensate those with $s(y) < 0$ through modifications to the tax schedule in the alternative environment. Proposition 2 states this formally using the scaled surplus function.

Proposition 2. Suppose Assumption 1 holds. For $\varepsilon > 0$, let $s_\varepsilon = \varepsilon s(y)$ denote the surplus offered by an $\varepsilon-$scaled version of the alternative environment. If $S > 0$, there exists $\tilde{\varepsilon} > 0$ such that for any $\varepsilon < \tilde{\varepsilon}$, there exists an augmentation to the tax schedule in the alternative environment that delivers surplus $s_\varepsilon^t(y)$ that is positive at all points along the income distribution, $s_\varepsilon^t(y) > 0$ for all y. Conversely, if $S < 0$, then no such $\tilde{\varepsilon}$ exists.

Proof. See Appendix A.3.

As in the Hicks (1940) experiment, the core insight of Kaldor (1939) is that one can compare the status quo and alternative environment using the Pareto principle, as opposed to relying on a particular social welfare function to make the comparison. I state this again formally in the following Corollary:

Corollary 2. For any set of (positive) social welfare weights, $\chi(y)$, the augmented status quo environment delivers greater social welfare than the alternative environment, $E[s_\varepsilon^t(y)\chi(y)] > E[s_\varepsilon(y)\chi(y)]$.

Proof. This follows from the fact that $s_\varepsilon^t(y) > s_\varepsilon(y)$ *for all y and that the weights* $\chi(y) > 0\ \forall y$.

In this sense, testing whether $S > 0$ provides a first-order approximation to searching for potential Pareto improvements as suggested by Kaldor (1939).

Summary

Table 1 summarizes the main results. When weighted surplus is negative, $S < 0$, the alternative environment is inefficient in the sense that a feasible modification to the tax schedule in the status quo environment can lead to a Pareto superior allocation to the alternative environment.

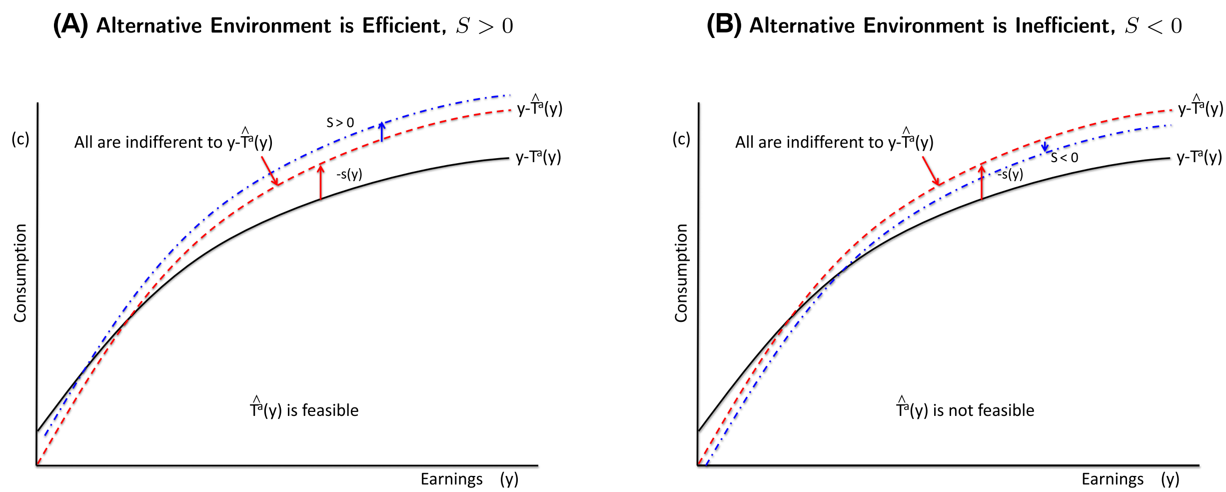

Fig. 4. Testing for (Kaldor) efficiency. *Notes*: This figure illustrates the test of efficiency in Kaldor (1939) that modifies the tax schedule in the alternative environment to attempt to find a Pareto improvement in the modified alternative environment relative to the status quo. The dashed red line presents the after-tax schedule that adds the surplus offered by the alternative environment to the tax schedule, $T(y) + s(y)$. To first order, everyone is indifferent between the status quo and the modified alternative environment illustrated by the dashed red line in Panels A and B. The dash-dot blue line then illustrates the after tax income curve that results from closing the government budget constraint. Panel A illustrates the case that the alternative environment is efficient, so that after modifying the tax schedule in the alternative environment there is a Pareto improvement relative to the status quo. Panel B illustrates the case where the alternative environment is inefficient, so that after taxing back the benefits of the alternative environment and closing the budget constraint everyone is worse off relative to the status quo.

In this sense, alternative environments for which $S < 0$ can be rejected by the logic of Hicks (1940) and Coate (2000). When weighted surplus is positive, $S > 0$, a modified alternative environment in which the winners compensate the losers through modifications to the tax schedule offers a Pareto superior allocation relative to the status quo. In this sense, the alternative can be preferred using the compensation principle in Kaldor (1939). In this sense, weighted surplus $S = E[s(y)g(y)]$ measures economic efficiency in the original spirit of Kaldor and Hicks.

4. Representing fiscal externalities using estimable parameters

As illustrated in Fig. 1, the marginal cost of providing a \$1 tax cut to those with earnings near y is given by $g(y) = 1 + FE(y)$, where $FE(y)$ is the impact of the behavioral response to the tax cut on government tax revenue. To estimate $FE(y)$, I build upon recent work by Jacobs et al. (2017) who provide an expression for $FE(y)$ as the sum of three components: a participation response, income effect, and substitution effect.

To captures these responses, let $\varepsilon^c(y)$ denote the average intensive margin compensated elasticity of earnings with respect to the marginal keep rate, $1 - \tau(y)$, for those earning $y(\theta) = y$:

$$\varepsilon^c(y) = E\left[\frac{1-\tau(y(\theta))}{y(\theta)}\frac{dy}{d(1-\tau)}|_{u=u(c,y;\theta)}|y(\theta) = y\right].$$

Let $\zeta(y)$ denote the average participation response in earnings to a percent increase consumption:

$$\zeta(y) = E\left[\frac{dy(\theta)}{dm}\frac{y(\theta)-T(y(\theta))}{y(\theta)}|y(\theta) = y\right]$$

Table 1
Comparisons of Hicks and Kaldor experiments.

	S>0	S<0
Hicks experiment: Possible to replicate s(y) using tax cut in status quo?	No	Yes
IKaldor experiment: Possible to modify alternative environment tax schedule to make everyone better relative to status quo?	Yes	No

And, let $\varepsilon^P(y)$ denote the average extensive margin (participation) elasticity with respect to net of tax earnings:

$$\varepsilon^P(y) = \frac{d[f(y)]}{d[y-T(y)]}\frac{y-T(y)}{f(y)}$$

where $f(y)$ is the density of income at y.

Appendix B extends the results in Jacobs et al. (2017) to allow for multi-dimensional heterogeneity and I provide formal assumptions under which which one can write the fiscal externality as the sum of these three types of responses:

$$FE(y) = \underbrace{-\varepsilon_c^P(y)\frac{T(y)-T(0)}{y-T(y)}}_{\text{Participation Effect}} \underbrace{-\zeta(y)\frac{\tau(y)}{1-\frac{T(y)}{y}}}_{\text{Income Effect}} \underbrace{-\varepsilon^c(y)\frac{\tau(y)}{1-\tau(y)}\alpha(y)}_{\text{Substitution Effect}} \quad (4)$$

where $\alpha(y) = -(1 + \frac{yf'(y)}{f(y)})$ is the local Pareto parameter of the income distribution.[14]

The first term in the fiscal externality arises from people entering the labor force to obtain the transfer. The participation elasticity, $\varepsilon^P(y)$, measures the size of this effect. The impact of this response on government revenue depends on the difference between the average taxes received at y, $T(y)$, and the taxes/transfers received from those out of the labor force, $T(0)$.

Second, the increased transfer may change the labor supply of those earning y due to an income effect. The size of this effect is measured by $\zeta(y)$. The impact of this response on government revenue depends on the marginal tax rate, $\tau(y)$.

Finally, people earning close to y may change their earnings towards y in order to get the transfer. The elasticity, $\varepsilon^c(y)$, measures how much people move their earnings towards y in response to the tax cut. The tax ratio, $\frac{\tau(y)}{1-\tau(y)}$, captures the impact of these responses on government revenue. However, the net impact is the sum of two types of

[14] As noted above, Eq. (4) is a generalization of the formula in Jacobs et al. (2017) to the case of multi-dimensional heterogeneity. Consistent with the intuition provided by Saez (2001), the relevant empirical elasticities in the case of potentially multi-dimensional heterogeneity are the population average elasticities conditional on income.

substitution responses. Some people will decrease their earnings towards y; others will increase their earnings towards y, as depicted by the blue arrows in Fig. 1. When $\tau(y) > 0$, the former effect increases tax revenue and the latter effect decreases tax revenue. The extent to which the losses outweigh the gains depends on the elasticity of the income distribution, $\frac{yf'(y)}{f(y)}$. When $\frac{yf'(y)}{f(y)} < -1$ (as is the case with the Pareto upper tails in the US income distribution), more people increase rather than decrease their taxable earnings. This means $\alpha(y) > 0$. Conversely, if $\frac{yf'(y)}{f(y)} > -1$ (e.g. if f is a uniform distribution so that $f'(y) = 0$), then more people decrease than increase their earnings so that $\alpha(y) < 0$. This increases the marginal cost of the tax cut. Importantly, this shows that even if elasticities and tax rates are constant, the shape of the income distribution plays a key role in determining the marginal cost of taxation at each income level.

5. Bounds on $g(y)$ in the U.S.

Before turning to an estimation of FE (y) in Eq. (4) (which follows in Section 6 below), this section first provides an estimation of the Pareto parameter of the income distribution, $\alpha(y)$. I use this to first place bounds on the shape of the weights, $g(y)$, without precise assumptions about the magnitude the behavioral elasticities.

I measure $\alpha(y)$ for each point of the income distribution using the universe of income tax returns from 2012.[15] Fig. 5 presents the mean value of $\alpha(y)$ at each quantile of the ordinary income distribution. The average $\alpha(y)$ reaches around 1.5 at the top of the income distribution, consistent with findings in previous literature focusing on top incomes (Diamond and Saez (2011) and Piketty and Saez (2013)). However, the key point on Fig. 5 is that $\alpha(y)$ exhibits considerable heterogeneity across the income distribution. It is negative below the 60th percentile of the income distribution, $\frac{yf'(y)}{f(y)} > -1$. This implies that the substitution effect increases the marginal cost of a tax cut (assuming a positive elasticity).[16] Conversely, it crosses zero around the 60th percentile, and is then positive.[17] This means that $\frac{yf'(y)}{f(y)} < -1$ for values of y above the 60th quantile. For those earning more than about \$43 K in ordinary income, the substitution effect reduces the cost of providing a tax cut. As long as $\tau(y) > 0$ and $\varepsilon^c(y) > 0$, the substitution effect, $-\varepsilon^c(y)\frac{\tau(y)}{1-\tau(y)}\alpha(y)$, in Eq. (4) is positive for incomes below \$43 K (60th quantile of 2012 ordinary income) and negative for incomes above \$43 K.

In addition to the substitution effect, it is also possible to put bounds on the natural shape of the impact of the participation effect on the government budget. For those with low incomes, the EITC offers transfers for those who enter the labor force; this renders $T(y) < 0$ so that those who enter the labor force in response to an increased tax cut actually increase the budgetary cost because they obtain additional transfers in the form of EITC benefits. In contrast, for higher values of y individuals contribute positive tax revenue so that $T(y) > 0$; thus any increase in labor force participation for those at higher income levels will result in a positive fiscal externality. This suggests the participation effect in Eq. (4) is also declining in y.

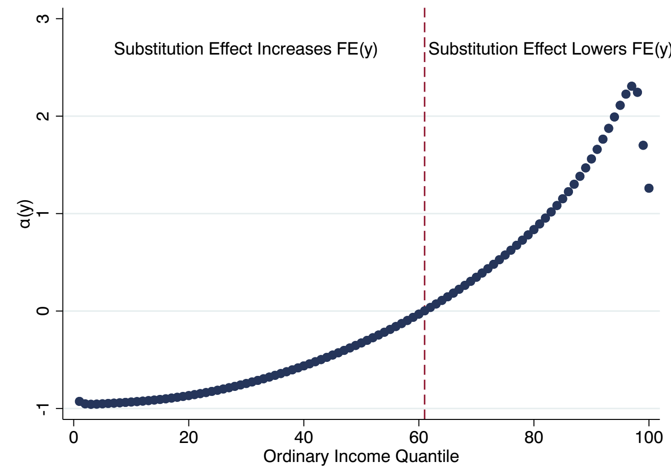

Fig. 5. Shape of the income distribution, $\alpha(y)$. *Notes*: This figure presents the estimates of the average $\alpha(y)$ for each quantile of the income distribution. This function is given by $\alpha(y) = -\left(1 + \frac{yf'(y)}{f(y)}\right)$, where $f(y)$ is the density of the income distribution. For values of y below the 60th quantile, $\alpha(y) < 0$ so that the substitution effect in Eq. (4) raises the marginal cost of taxation. In contrast, for values of y above the 61st quantile, $\alpha(y) > 0$ so that the substitution effect lowers the marginal cost of taxation.

Lastly, most empirical works suggests income effects effects are either small (Gruber and Saez (2002); Saez et al., 2012) or declining in income (Cesarini et al. (2015)). As a result, one has a natural bound on the shape of the weights, $g(y)$: the welfare weights put greater weight on those with lower incomes (i.e. below \$43 K) than those with higher incomes (i.e. above \$43 K). This means that it costs the government more than \$1 to provide an additional dollar of benefits to a person with a low incomes but less than \$1 to provide an additional \$1 of benefits to a person with high incomes.

6. Using elasticities to quantify $g(y)$ in the U.S.

Point estimates of $g(y)$ require estimates of the behavioral responses to taxation. For those subject to the EITC, I draw upon Chetty et al. (2013) who calculate elasticities of 0.31 in the phase-in region (income below \$9560) and 0.14 in the phase-out region (income between \$22,870 and \$43,210). Using the income tax return data, I assign these elasticities to EITC filers in these regions of the income distribution. Second, for filers subject to the top marginal income tax rate, I assign a compensated elasticity of 0.3. This is consistent with the midpoint of estimates estimated from previous literature studying the behavioral response to changes in the top marginal income tax rate (Saez et al. (2012)). For those not on EITC and not subject to the top marginal income tax rate, I assign a compensated elasticity of 0.3, consistent with Chetty (2012) who shows such an estimate can rationalize the large literature on the response to taxation. I assess the robustness to alternative elasticities such as 0.1 and 0.5.

In addition to these intensive margin responses, there is also significant evidence of extensive margin behavioral responses, especially for those subject to the EITC. This literature suggests EITC expansions are roughly 9% more costly to the government due to extensive margin behavioral responses.[18] Therefore, I assume the participation effect in eq. (4) is equal to 0.09 for income groups subject to the EITC. Above the EITC range, there is mixed evidence of participation responses to taxation. Liebman and Saez (2006) find no statistically significant impact of tax changes on women's labor supply of women married to higher-income men. Indeed, higher tax rates can reduce participation from a

[15] Formally, I construct this by separately estimating $\alpha(y)$ for each tax schedule using the information in the tax returns on filing status and other determinants of the tax schedule. As noted in the Appendix, throughout I estimate $g(y)$ using a method that correctly accounts for the heterogeneity in tax schedules faced by those at the same level of income. The details of this procedure are provided in Appendix E

[16] This is consistent with the findings of Werning (2007) who estimates the marginal cost of taxation using the SOI public use file.

[17] The shape is non-monotonic at the top of the distribution, which reflects the fact that the US income distribution has roughly a log-normal shape throughout much of the distribution and transitions into a Pareto tail in the top of the distribution. Log-normality would have meant $\alpha(y) \to \infty$; but for Pareto tails $\alpha(y)$ converges to a constant.

[18] See Hotz and Scholz (2003) for a summary of elasticities and Hendren (2016) for the 9% calculation.

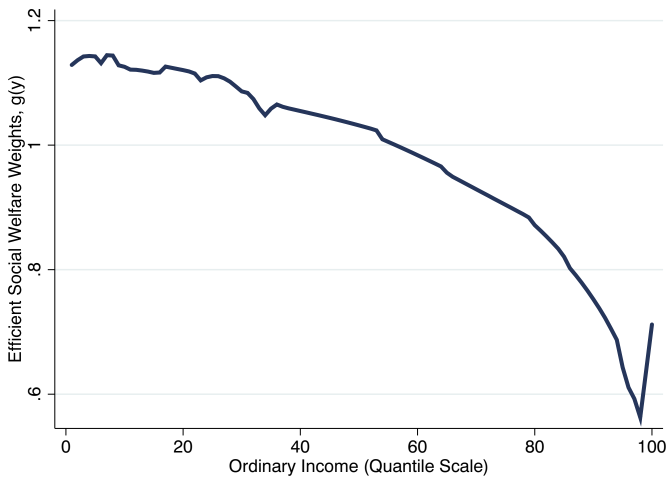

Fig. 6. Inverse-Optimum Welfare Weights, $g(y)$. *Notes*: This figure presents the baseline estimates of the weights, $g(y)$, estimated using Eq. (4) for each quantile of the income distribution.

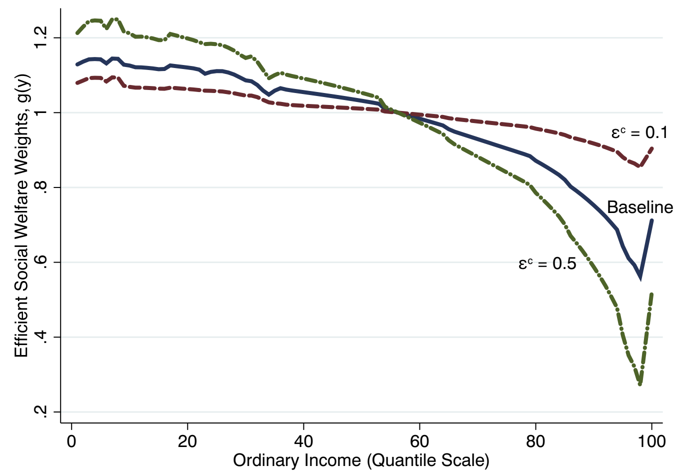

Fig. 7. Robustness to alternative elasticities. *Notes*: This figure presents the baseline specification for the inverse-optimum welfare weights alongside with estimates under alternative constant compensated elasticity scenarios of $\varepsilon^c(y) = 0.1$ and $\varepsilon^c(y) = 0.5$.

price effect but increase participation due to an income effect. As a result, I assume a zero participation elasticity for those not subject to the EITC.

Lastly, I assume away intensive margin income effects, consistent with a large literature suggesting such effects are small (Gruber and Saez (2002); Saez et al. (2012)). Cesarini et al. (2015) find evidence of income effects using Swedish lotteries; however a large portion of these effects are driven by extensive margin responses and arguably already captured by the EITC responses measured above.[19]

Results

I use Eq. (4) to combine the estimates of the shape of the income distribution, marginal tax rates, and elasticity calibrations, which generates an estimate of $FE(y)$ for each filer. I then bin the income distribution into 100 quantile bins and construct the mean fiscal externality, $FE(y)$, for each quantile of income. The inverse-optimal weight at each income quantile is then given by $g(y) = 1 + FE(y)$. Fig. 6 presents the resulting estimates for $g(y)$. Fig. 7 presents the results for the alternative calibrations of the compensated elasticity of $\varepsilon^c = 0.1$ and $\varepsilon^c = 0.5$.

The weights have several key features. First, consistent with the bounds shown in the previous section, the results suggest it is efficient to place higher weight on surplus to the poor than to the rich. Under the baseline specification, these weights fall from around 1.15 for those at the bottom of the income distribution to 0.65 for those at the top. Transferring \$1 from the top of the distribution can generate around $0.65/1.15 = \$0.57$ of welfare to someone at the bottom of the distribution. Conversely, transferring \$1 from the bottom of the income distribution can generate around $1.15/0.65 = \$1.77$ of welfare to the those at the top of the income distribution.

Second, although the weights place more weight on low versus high income individuals, the weights never differ by more than a factor of 2. In other words, $|\frac{g(y)}{g(y')}| < 2$ for all y and y'. This means that it is not efficient to discount surplus more than 50%, regardless of where it falls in the income distribution. For example, the consumer surplus standard in merger analysis (which gives no weight to producer surplus) would still not be efficient even after accounting for the distortionary cost of taxation.

Third, while the weights generally decline in income, there is an increase in the top 1%. This reversal is highly statistically significant as seen by the drop in $\alpha(y)$ in Fig. 5. Statistically, this drop reflects the transition of the income distribution from a 'log-normal' shape (for which $\alpha(y)$ would be increasing as y increases) to a Pareto distribution in which $\alpha(y)$ converges to a constant. Because $\alpha(y)$ rises above 1.5 and then falls in the top regions of the income distribution, this means that the marginal cost of taxation is lower for the upper middle class than the top 1% (assuming a constant compensated elasticity of taxation). However, Fig. 7 illustrates that this non-monotonicity is not robust to plausible assumptions about how elasticities might change across the income distribution. In particular, if the elasticity moves from 0.3 to 0.5 as one goes from the top 2% to the top 1%, the weights would again be monotonically declining in income.

Fourth, all the weights are positive, $g(y) > 0$ for all y for the baseline and alternative specifications. This means that it is always costly to provide a tax cut. This implements a Pareto efficiency test suggested by Werning (2007), and suggests there are no Pareto improvements solely from modifying the tax schedule.

Lastly, as foreshadowed by the bounding exercise in the previous Section, there is a similarity between the estimates of $\alpha(y)$ in Fig. 5 and the shape of the weights, $g(y)$. Higher elasticities, $\varepsilon^c(y)$, increase the difference between the weights on the low- versus high-income individuals. But, they do not affect the general conclusion that $g(y) > 1$ for those with low incomes and $g(y) < 1$ for those with high incomes.

7. Applications: comparison of income distributions

> [*Using transfers*], "*it is always possible for the Government to ensure that the previous income-distribution should be maintained intact*" (Kaldor (1939)).

Kaldor and Hicks' original motivation was the comparison of different distributions of endowments. Motivated by this classic comparison, I use the weights, $g(y)$, to compare distributions of income. I begin with an analysis of changes in the U.S. income distribution over time; I then explore cross-country differences in income distributions.

To compare income distributions, one needs to define a conceptual experiment that clarifies where an individual in one distribution would fall in the alternative distribution. This experiment then defines

[19] Nonetheless, Appendix F reports the robustness of the results to an alternative specification that incorporates income effects assuming that the estimates from Cesarini et al. (2015) are entirely along the intensive margin and correspond to an elasticity of $\zeta = 0.15$. As shown in Appendix Fig. 3, income effects tend to increase the marginal cost of taxation at all income levels; but in contrast to the compensated elasticity they do not affect the relative difference in the weights to low versus high income individuals.

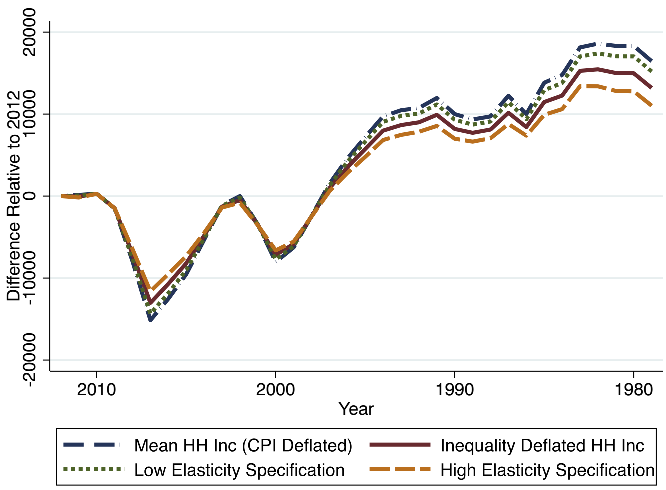

Fig. 8. Raw and Deflated Household Income Change Relative to 2012. *Notes*: This figure presents the un-weighted growth in incomes (relative to 2012) and the distributionally-adjusted growth in incomes using the low elasticity ($\varepsilon = 0.1$), baseline elasticity ($\varepsilon = 0.3$), and high elasticity ($\varepsilon = 0.5$) specifications.

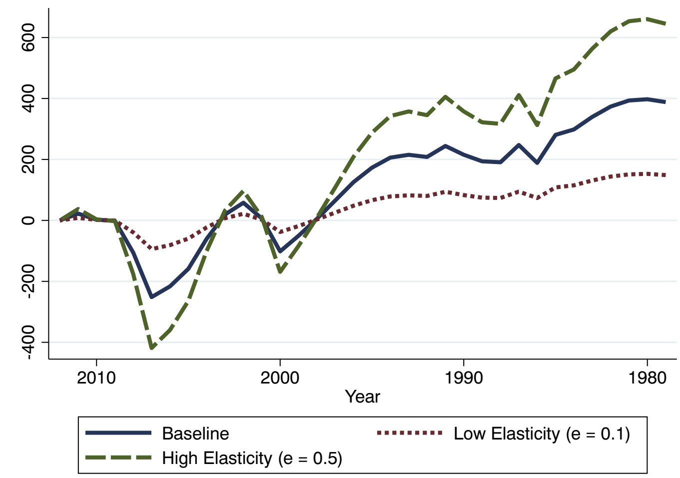

Fig. 9. Social cost of increased income inequality. *Notes*: This figure presents a measure of the social cost of income inequality, which is defined as the difference between un-weighted growth in incomes since 2012 and distributionally-adjusted growth in incomes, multiplied by the number of households in the US of 117.6 M. The results are presented using the low elasticity ($\varepsilon = 0.1$), baseline elasticity ($\varepsilon = 0.3$), and high elasticity ($\varepsilon = 0.5$) specifications.

the surplus function, $s(y)$, that can be used to compare the distributions. In general, one could imagine comparisons between two income distributions where people who are at the top of the distribution stay at the top of the distribution in the alternative environment; conversely one could imagine an experiment where people at the top switch with those at the bottom. For simplicity, I consider the experiment here in which each person's relative position in the income distribution is maintained in tact (i.e. someone with median income in the status quo distribution is at the median in the alternative distribution).

Finally, I use the weights, $g(y)$, estimated in the US. This means that formally I will implement a first order test of the Hicks (1940) and Coate (2000) test for efficiency outlined in Section 3.1. For comparisons over time, I ask how much better off the US is in 2012 relative to earlier years if it attempted to replicate the shape of the income distribution in those earlier years so that the growth was spread equally throughout the distribution. For cross country comparisons, this asks how much better or worse off the US income distribution is relative to other countries after using the tax schedule to replicate their income distribution. To the extent to which the marginal cost of taxation is the same at each quantile of the distribution in other countries, this also implements the Kaldor (1939) test for efficiency.[20]

7.1. Income growth in the U.S.

It is well-known that income inequality in the U.S. has increased in recent decades, especially at the top of the distribution (Piketty and Saez (2003)). Appendix Fig. 1 plots several quantiles of the household after-tax income distribution over time using data from the Congressional Budget Office (CBO) from 1979 to 2009.[21] As is well-known, incomes have increased significantly in the top portions of the income distribution, especially the top 20% and top 1%; in contrast, income for the bottom 80% has experienced smaller growth.

Here, I use the inverse-optimum welfare weights, $g(y)$, to calculate how much richer all points of the income distribution would be relative to a given previous year if the tax schedule were augmented in order to remove any changes in income inequality over time. Let $Q_0(\alpha)$ denote the α-quantile of the 2012 income distribution; let $Q_t(\alpha)$ denote the α-quantile of an alternative income distribution in year t. I define the weighted surplus in household income by

$$S_t = \int_0^1 [Q_0(\alpha) - Q_t(\alpha)] g^H(Q_0(\alpha)) d\alpha \quad (5)$$

where $g^H(y)$ are the inverse-optimum weights. Intuitively, S_t is the first-order approximation to the amount by which the U.S. would be richer in 2012 relative to year t if the 2012 income tax schedule were augmented in to hold constant the changes to the income distribution relative to year t. All incomes are in units of 2012 income using the CPI-U deflator.

Fig. 8 reports the change in mean household income (dashed blue line), along with the weighted surplus under the baseline specification and two alternative elasticity specifications. Mean household income has increased by roughly \$18,300 relative to 1979, but if these benefits were redistributed equally across the population, growth would have has increased \$15,000 under the baseline specification (\$13,000 and \$17 K under the high and low elasticity specifications, respectively). From a normative perspective, this lowers the overall growth rate of the U.S. economy by roughly 15–20%: if the U.S. were to make a tax adjustment so that everyone shared equally in the after tax earnings increases, roughly 15–20% of the growth since 1979 would be evaporated.

Fig. 9 provides an estimate of the social cost of increased income inequality. To do so, I multiply the per-household social cost by the total number of households in the U.S.[22] This suggests the social cost of increased income inequality since 1979 is roughly \$400B. From an equivalent variation perspective, undoing the increased inequality would cost roughly \$400B; from a compensating variation perspective, if the U.S. had not experienced the increased inequality, it could have replicated the social surplus provided by the 2012 after tax income distribution even if aggregate economic growth were \$400B less than actually

[20] I leave an analysis of the difference in weights across countries or over time to the growing existing and future body of work estimating these weights.

[21] The data is constructed using Table 7 from CBO publication 43,373. I take market income minus federal taxes to construct after-tax income shares across the population. To account for the fact that government spending may have value, I assign net tax collection back to each household in proportion to their after-tax income. This assumes each individuals' willingness to pay for government expenditure is proportional to after-tax income. The CBO also reports an "after-tax" measure of income that includes government transfers. Unfortunately, the bottom portion of the income distribution for these transfers disproportionately falls on the non-working elderly, through social security and Medicare payments. Since these would be affected by modifications to the nonlinear income tax schedule, I do not use this measure of income.

[22] The census reports 117.6 M households in 2009, with an annual increase over the years 2006–2009 of roughly 500 households per year, implying roughly 119 M households in 2012.

occurred. These numbers depend on the behavioral responses to taxation – if one believes behavioral responses to taxes are larger (e.g. a compensated elasticity of 0.5), then the social cost of increased income inequality is in excess of $600B.

To be sure, the comparison of the income distribution in 2012 to the income distribution in 1979 is perhaps not best thought of as a "marginal" policy comparison. To that aim, the most robust conclusion that can be drawn from the analysis above is the following: if the distribution of economic growth continued from today to follow the average trend in the US since 1979, then unweighted measures of economic growth will over-state the growth in societal well-being by roughly 15–20%. This 15–20% statistic holds exactly when considering small amounts of economic growth (i.e. short time windows), but as discussed below in Section 8.1, it could differ when considering larger differences in the income distribution if the marginal cost of taxation changes as one modifies the tax schedule. An important direction for future work is understanding how changes in the tax schedule lead to changes in the inverse-optimum welfare weights, which could then be used to adjust for these second-order effects.

7.2. Comparisons of income distributions: cross-country analysis

It is often noted that the U.S. has a higher degree of income inequality than many other countries of similar income per capita levels. In this subsection, I use the inverse-optimum welfare weights to ask how much richer or poorer the U.S. would be relative to these countries if it attempted to replicate their income distributions using modifications to the tax schedule.

The weighted surplus associated with moving from the status quo income distribution to the income distribution in country a is given by

$$S_a^{ID} = \int_0^1 [Q_a(\alpha) - Q_0(\alpha)] g^H(Q_0(\alpha)) d\alpha \tag{6}$$

I form estimates of $Q_a(\alpha)$ using data from the World Bank Development Indicators and UN World Income Inequality Database. These sources aggregate household survey data from various countries and to provide measures of the shape of the income distribution.

Fig. 10 plots deflated surplus against the GNI per capita of each country within $10,000 of the U.S. GNI per capita. The dots represent the estimates for the baseline specification and the brackets plot the estimates for the low and high elasticity specifications.

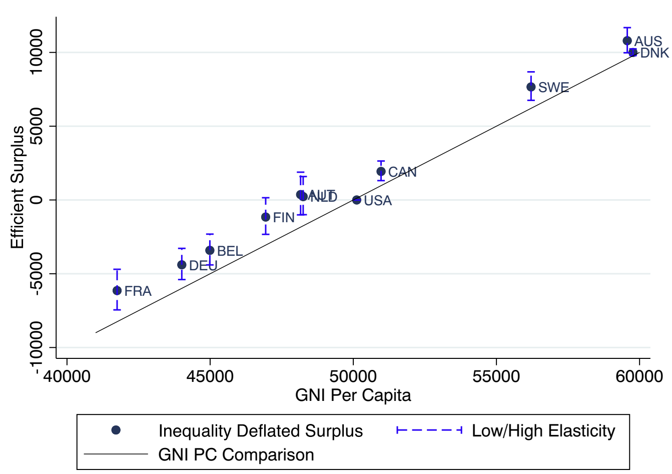

Fig. 10. Comparisons of income distributions across countries. *Notes*: This figure plots weighted surplus and GNI per capita for a selection of countries with gross national incomes (GNI) near that of the US. For each country, the weighted surplus (defined in Eq. (6)) is presented for the baseline elasticity specification against GNI per capita on the horizontal axis; vertical bars representing the high and low elasticity specifications. If all countries had the same degree of inequality, then all countries would align on the 45 degree line. The fact that other countries lie above this 45 degree line reflects the greater degree of income inequality in the U.S. relative to these countries.

The results suggest that a couple of cross-country comparisons based on mean incomes are reversed when using the inverse-optimum weights to control for differences in inequality. The U.S. is richer in mean per capita terms than Austria (AUT) and New Zealand (NLD) by roughly $2000. But despite it's higher income level, if the U.S. were to try to provide the distribution of purchasing power offered by these countries through modifications to the tax schedule, each point of the income distribution would be made worse off relative to these countries under the baseline elasticity specification. Under the high elasticity specification, it would be efficient to take Finland's income distribution over the US's income distribution, even though it has $3180 less in per capita national income.

8. Discussion of limitations of the approach

8.1. Non-marginal comparisons

The formal results above show that weighting surplus by the inverse-optimum welfare weights search for potential Pareto improvements for small surplus comparisons. In practice however, many comparisons of interest are likely not best thought of as "small". In these instances, there are two potential concerns that can arise.

First, for non-marginal transfers, the marginal cost of the first dollar of the transfers may not equal the marginal cost of the last dollar of the transfers. In this case, $E[s(y)g(y)]$ would not accurately measure the revenue that the government is able to one would prefer to use the weight that measures the average cost of providing $s(y)$ to each level of income.

Second, if the alternative environment is sufficiently distinct from the status quo, then an individuals' willingness to pay will depend on whether it is paid out of income in the status quo or alternative environment. The definition of $s(\theta)$ above is an "equivalent variation" definition of willingness to pay because it imagines this amount being paid out of income in the status quo. Another method for measuring willingness to pay would be to consider a "compensating variation" definition, which would imagine a willingness to pay out of income in the alternative environment. To first order, these two definitions of willingness to pay are always equivalent. But, they generally differ in non-marginal comparisons.

While compensating and equivalent variation measures of surplus can differ in general, they are identical when comparing environments where the difference across environments is one's income. An individual is always willing to pay $10 to receive $10 of additional income – this is true whether one conceptualizes willingness to pay as an amount of income needed to give someone in the status quo world to make them indifferent to receiving $10 (equivalent variation), or as the amount of income one can take away in the alternative environment to make them indifferent to not receiving the additional income (compensating variation). This means this concern does not apply to the results in Section 7, but could be relevant in other comparisons.

8.2. General equilibrium effects

Second, the approach assumes that tax changes have no general equilibrium or spillover effects. Targeting a $1 tax cut to those earning near $y is assumed to have a willingness to pay of $1 for the beneficiaries of the tax cut. But, if their wages change in response to the tax cut, their willingness to pay may differ from $1. Indeed, with spillovers and general equilibrium effects, the benefits of the tax cut may extend beyond those who are the direct target of the tax cut. But while taxation is not allowed to have GE effects, the approach does allow GE effects to drive the valuation of the alternative environment, $s(y)$. For example, the alternative environment could be a policy that makes more land available for agriculture, which in turn lowers food prices. One can still generate individuals' willingness to pay for this alternative environment, $s(y)$, and use the inverse-optimum welfare weights to ask whether this policy is efficient. In this sense, the weights, $g(y)$, are valid even if the policy change or alternative environment has GE effects; but it has ruled out the case where changes in the tax schedule, $T(y)$, have GE effects. Recent work by Tsyvinski and Werquin (2018) generalizes the approach provided here to allow for taxation with GE effects.

8.3. Heterogeneity in $s(\theta)$ conditional on y

Third, alternative environments may generate a willingness to pay that is heterogeneous conditional on income. In this case, Pareto comparisons are more difficult. Appendix D shows that to test for Hicks efficiency, one needs to construct the maximum willingness to pay at each income level, $\overline{s}(y)$, and test whether $E[\overline{s}(y)g(y)]>0$. If it is negative, then it would be feasible for the government to replicate the surplus offered by the alternative environment and make everyone better off. Intuitively, the government can feasibly provide a tax cut that covers even the maximal willingness to pay at each income level, $\overline{s}(y)$. In this sense, the alternative environment would be inefficient. Conversely, to test for Kaldor efficiency, one needs to construct the minimum willingness to pay at each income level, $\underline{s}(y)$, and test whether $E[\underline{s}(y)g(y)]>0$. If it is positive, then it would be feasible for the government to redistribute income in the alternative environment so that everyone prefers the modified alternative environment relative to the status quo.[23] Often, one might find that $E[\underline{s}(y)g(y)]<0$ and $E[\overline{s}(y)g(y)]>0$. In this instance, the alternative environment cannot not be Pareto-ranked relative to the status quo. Nonetheless, the weights, $g(y)$, continue to be the key component required to measure $E[\underline{s}(y)g(y)]$ and $E[\overline{s}(y)g(y)]$ that facilitates the search for these Pareto comparisons.

8.4. The weights, $g(y)$ are not structural

Lastly, as noted above, the weights $g(y)$ are not structural parameters. They are endogenous to the economic environment. In addition to weights changing as one implements transfers, there is also no reason to expect weights identified in one setting or country to readily translate to another setting. For example, Bourguignon and Spadaro (2012) estimate weights for France that are close to zero at the top of the income distribution, suggesting that reductions in tax rates nearly pay for themselves, $g(y) \approx 0$. Adopting those results means that measuring economic efficiency in the French context would place lower weight on surplus accruing to the rich than in the US. Measuring economic efficiency requires adjusting for the cost of taxation, which can differ across settings.

9. Conclusion

In their original work, Kaldor and Hicks hoped to provide a method to avoid the inherent subjectivity involved in resolving interpersonal comparisons. This paper provides a straightforward approach to implement their classic efficiency experiments in a manner that accounts for the distortionary cost of taxation. Weighting surplus using inverse-optimum welfare weights measures the economic efficiency as envisioned by Kaldor and Hicks. Estimates for the US suggest that redistribution from rich to poor is more costly than from poor to rich. Thus, it is efficient to place greater weight on the poor than on the rich. If weighted surplus is positive, modifying the tax schedule in the alternative environment can make everyone better off relative to the status quo. This means that for any social welfare weights, the modified alternative environment would be preferred to the status quo. Conversely, if weighted surplus is negative, everyone can be made better off by modifying the tax schedule in the status quo relative to adopting the alternative environment. In this sense, weighted surplus measures economic efficiency as envisioned by Kaldor and Hicks. It generates a preference over alternative environments without appealing to social welfare weights.

There are many important directions for future work, including incorporating the general equilibrium effects of taxation (as in ongoing work by Tsyvinski and Werquin (2018)). Additionally, one could extend the analysis here to construct weights that involve redistribution not just through the tax schedule but also via other means, such as health insurance subsidies or other policies. By expanding the dimensionality of the weights, it could help deal with settings where surplus varies conditional on income. Appealing to the Pareto principle in the Kaldor and Hicks' experiments requires implementing the transfers that they envision. Future work could discuss the implications of political economy or other constraints that might prevent such transfers in practice. Lastly, the approach developed here is valid to first order, and it would be especially valuable to extend the analysis to non-marginal comparisons.

[23] Appendix D provides formal statements and proofs of these claims.

References

Atkinson, A. B. and J. E. Stiglitz (1976). The design of the tax structure: Direct versus indirect taxation. *Journal of Public Economics 6* (1-2). versus indirect taxation. *Journal of Public Economics 6* (1-2).

Bargain, O., M. Dolls, D. Neumann, A. Peichl, and S. Siegloch (2011). Tax-benefit systems in europe and the us: between equity and efficiency. *CESifo Working Paper No. 3534* .

Bargain, O., Dolls, M., Neumann, D., Peichl, A., Siegloch, S., 2014. Comparing inequality aversion across countries when labor supply responses differ. International Tax and Public Finance 21, 845–873.

Bloom, H., Orr, L.L., Bell, S.H., Cave, G., Doolittle, F., Lin, W., Boss, J.M., 1997. The benefits and costs of jtpa title ii-a programs: Key findings from the national job training partnership act study. The Journal of Human Resources 32 (3), 549–576.

Blundell, R., Brewer, M., Haan, P., Shephard, A., 2009. Optimal income taxation of lone mothers: An empirical comparison of the uk and germany. The Economic Journal 119, 101–121.

Bourguignon, F., Spadaro, A., 2012. Tax-benefit revealed social preferences. Journal of Economic Inequality 10, 75–108.

Cesarini, D., Lindqvist, E., Notowidigdo, M.J., Östling, R., 2015. The effect of wealth on individual and household labor supply: evidence from swedish lotteries. Technical report, National Bureau of Economic Research.

Chetty, R., 2012. Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica 80 (3), 969–1018.

Chetty, R., Friedman, J., Saez, E., 2013. Using differences in knowledge across neighborhoods to uncover the impacts of eitc on earnings. *American Economic Review*. 103 (7), 2683–2721 (December).

Chetty, R., Hendren, N., Kline, P., Saez, E., 2014. Where is the land of opportunity: The geography of intergenerational mobility in the united states. The Quarterly Journal of Economics 129 (4), 1553–1623.

Christiansen, V., 1977. The theoretical basis for deriving distributive weights to be used in cost- benefit analysis. Technical report, Memorandum from the Institute of Economics. University of Oslo 25 April.

Christiansen, V. (1981, 07). Evaluation of Public Projects under Optimal Taxation. The Review of Economic Studies 48 (3), 447–457.

Christiansen, V. and E. S. Jansen (1978). Implicit social preferences in the norwegian system of indirect taxation. Journal of Public Economics *10* (2), 217 – 245.

Coate, S., 2000. An efficiency approach to the evaluation of policy changes. Economic Journal 110 (463), 437–455.

Diamond, P. A. and E. Saez (2011). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives* 25 (4), 165–90.

Eissa, N., Kleven, H.J., Kreiner, C.T., 2008. Evaluation of four tax reforms in the united states: Labor supply and welfare effects for single mothers. Journal of Public Economics 92 (3-4), 795–816.

Gruber, J., Saez, E., 2002, April. The elasticity of taxable income: evidence and implications. Journal of Public Economics 84 (1), 1–32.

Hendren, N., 2016. The policy elasticity. Tax Policy and the Economy 30.

Hendren, N., Sprung-Keyser, B.D., 2019. A unified welfare analysis of government policies. Technical report, National Bureau of Economic Research.

Hicks, J.R., 1939. The foundations of welfare economics. Economic Journal 49, 696–712.

Hicks, J.R., 1940. The valuation of social income. Economica 7 (26), 105–124.

Hotz, J., Scholz, K., 2003. The earned income tax credit. In: Moffit, R.A. (Ed.), Tax Policy and the Economy. University of Chicago Press.

Hoynes, H.W., Schanzenbach, D.W., 2012. Work incentives and the food stamp program. Journal of Public Economics 96 (1-2), 151–162.

Hylland, A., Zeckhauser, R., 1979. Distributional objectives should affect taxes but not program choice or design. The Scandinavian Journal of Economics 81 (2), 264–284.

Immervoll, H., Kleven, H.J., Kreiner, C.T., Saez, E., 2007. Welfare reform in european countries: A microsimulation analysis. The Economic Journal 117, 1–44.

Immervoll, H., Kleven, H.J., Kreiner, C.T., Verdelin, N., 2011. Optimal tax and transfer programs for couples with extensive labor supply responses. Journal of Public Economics 95, 1485–1500.

Jacob, B.A., Ludwig, J., 2012. The effects of housing assistance on labor supply: Evidence from a voucher lottery. American Economic Review 102 (1), 272–304.

Jacobs, B.A., Jongen, E.L.W., Zoutman, F.T., 2017. Revealed social preferences of dutch political parties. *Journal of Public Economics* 165.

Kaldor, N., 1939. Welfare propositions of economics and interpersonal comparisons of utility. The Economic Journal 49 (195), 549–552.

Kaplow, L., 1996. The optimal supply of public goods and the distortionary cost of taxation. National Tax Journal 49, 513–534.

Kaplow, L., 2004. On the (ir)relevance of the distribution and labor supply distortion to government policy. Journal of Economic Perspectives 18, 159–175.

Kaplow, L., 2008. The Theory of Taxation and Public Economics. Princeton University Press.

Kleven, H.J., Kreiner, C.T., 2006. The marginal cost of public funds: Hours of work versus labor force participation. Journal of Public Economics 1955–1973.

Liebman, J., Saez, E., 2006. Earnings responses to increases in payroll taxes Working Paper.

Lockwood, B.B., Weinzierl, M., 2016. The evolution of revealed social preferences in the united states and the costs of unequal growth and recessions. Journal of Monetary Economics 77, 30–47.

Mayshar, J., 1990. On measures of excess burden and their applications. Journal of Public Economics 43 (3), 263–289.

Mirrlees, J.A., 1971. An exploration into the theory of optimal income taxation. The Review of Economic Studies 38 (2), 175–208.

Piketty, T., Saez, E., 2003. Income inequality in the united states, 1913-1998. Quarterly Journal of Economics 118 (1), 1–41.

Piketty, T., Saez, E., 2013. Optimal labor income taxation. Handbook of Public Economics 5, 391.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68, 205–229.

Saez, E., Stantcheva, S., 2016. Generalized social marginal welfare weights for optimal tax theory. American Economic Review 106 (1), 24–45.

Saez, E., Slemrod, J., Giertz, S.H., 2012. The elasticity of taxable income with respect to marginal tax rates: A critical review. Journal of Economic Literature 50 (1), 3–50.

Schlee, E.E., 2013. Radner's cost-benefit analysis in the small: An equivalence result Working Paper.

Slemrod, J., Yitzhaki, S., 1996. The social cost of taxation and the marginal cost of funds. International Monetary Fund Staff Papers 43 (1), 172–198.

Slemrod, J., Yitzhaki, S., 2001. Integrating expenditure and tax decisions: The marginal cost of funds and the marginal benefit of projects. *National Tax Journal*. 54 (2), 189–202.

Tsyvinski, A. and N. Werquin (2018). Generalized compensation principle. Technical report, National Bureau of Economic Research, Working Paper 23509.

Werning, I., 2007. Pareto efficient income taxation MIT Working Paper.

Zoutman, F.T., Jacobs, B., Jongen, E.L.W., 2013. Optimal redistributive taxes and redistributive preferences in the netherlands Mimeo.