Pareto-improving tax reforms and the Earned Income Tax Credit*

Felix J. Bierbrauer† Pierre C. Boyer‡ Emanuel Hansen§

August 26, 2021

Abstract

This paper provides necessary and sufficient conditions for the existence of Pareto-improving tax reforms. The conditions can be expressed as sufficient statistics and have a wide range of potential applications in public finance. We discuss one such application in detail: the introduction of the Earned Income Tax Credit (EITC) in the US. We find that the EITC can be viewed as a response to an inefficiency in the tax and transfer system prevailing at the time. This adds a new perspective to the literature on why the EITC is a good idea, emphasizing Pareto improvements rather than equity-efficiency trade-offs.

Keywords: Tax reforms; Non-linear income taxation; Optimal taxation; Earned Income Tax Credits; Pareto Efficiency.

JEL classification: C72; D72; D82; H21.

---

*We thank Thomas Gaube, Ulrich Glogowsky, Bas Jacobs, Laurence Jacquet, Etienne Lehmann, Jonas Löbbing, Andreas Peichl, Dominik Sachs, Florian Scheuer, Laurent Simula, Alain Trannoy, Nicolas Werquin and seminar audiences in Brussels, Marseille, Lyon, Cologne, Munich, Paris, Florence, and Osnabrück for helpful comments. Susanna Bolz and Marvin Immesberger provided excellent research assistance. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2126/1 – 390838866, the Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047), ANR-19-CE41-0011-01 and ANR-20-CE41-0013-01 are gratefully acknowledged.

†CMR - Center for Macroeconomic Research, University of Cologne, Albert-Magnus Platz, 50923 Köln, Germany. E-mail: bierbrauer@wiso.uni-koeln.de

‡CREST, École polytechnique, Institut Polytechnique de Paris, 5, avenue Henry Le Chatelier, 91120 Palaiseau, France. E-mail: pierre.boyer@polytechnique.edu

§CMR - Center for Macroeconomic Research, University of Cologne, Albert-Magnus Platz, 50923 Köln, Germany. E-mail: hansen@wiso.uni-koeln.de

1 Introduction

This paper presents new results on Pareto-efficient income taxation. Specifically, it provides necessary and sufficient conditions for the existence of a Pareto-improving reform direction. The analysis is based on a general framework that nests prominent models of taxation as special cases. We foresee a range of potential applications in public finance that combine our characterization with sufficient statistics for the revenue implications of tax reforms.

We discuss one such application in detail, the introduction of the Earned Income Tax Credit (EITC) in the US in the mid-1970s. For this application, we derive sufficient statistics from a model with fixed costs of labor market participation and variable costs of productive effort.[1] We find that the introduction of the EITC can be viewed as a response to an inefficiency in the tax and transfer system prevailing at the time. The judgment that the introduction of an EITC was a "good idea" is remarkably robust: it holds for any Paretian welfare function and for all empirically plausible values of labor supply elasticities at the intensive and the extensive margin.

A theory of Pareto-improving tax reforms. Our theoretical analysis is motivated by two observations: first, past reforms of the EITC typically involved *two brackets*, a phase-in range with lower marginal tax rates and a phase-out range with higher tax rates. Second, an observation on the typical thought experiment in the literature that uses perturbation methods for a characterization of optimal tax systems: it analyzes the welfare implications of lowering or raising the marginal tax rates in *one bracket*.[2]

These observations raise the question whether reforms with two brackets can do "more" than reforms involving a single bracket. Suppose that a given tax system

[1] This framework is prominent in the literature that studies the EITC from an optimal tax perspective – see Saez (2002), Choné and Laroque (2011), Jacquet, Lehmann and Van der Linden (2013), or Hansen (2021).

[2] See Piketty (1997), Saez (2001), Golosov, Tsyvinski and Werquin (2014), or Jacquet and Lehmann (2021).

is "one-bracket-efficient" in the sense that there does not exist a Pareto-improving one-bracket reform. Can there be reforms with two brackets that make everyone better off? We show that the answer is "yes", i.e., reforms involving two brackets can achieve more than reforms with one bracket.

This leads to the next question. Suppose that the scope for Pareto-improving two-bracket reforms has been exhausted. Can there be reforms with three or even more brackets that make everyone better off? We show that the answer is "no", i.e., if there is no Pareto-improving reform involving one or two brackets, then there is no Pareto-improving element in the set of continuous reform directions.

These findings are derived from a generic static model of taxation: Individuals derive utility from consumption and the generation of income requires costly effort. They face a budget constraint that is shaped by a predetermined non-linear tax-transfer system. We consider reforms of this tax system so that marginal tax rates are changed simultaneously in an arbitrary number of income brackets. Also, there is full flexibility in terms of locating those brackets in the range of possible incomes. We then focus on the limit case of small reforms, involving marginal changes of tax rates over finitely many brackets of infinitesimal length. The interpretation is that we consider directions for reform in a neighborhood of a given status quo.

Making use of the theory. Our results provide guidance for the design of tax systems. There are two broad insights: *"Two is more than one!"*, one should not miss the additional opportunities that come with two-bracket reforms. *"Two is enough!"*, one does not miss reform opportunities by focusing on reforms with one or two brackets.

How can one use these insights? More specifically, how can one figure out whether a given tax and transfer system admits a Pareto-improving reform? Our analysis yields a test function that gives, for each income level y, the revenue implication $\mathcal{R}(y)$ of a small one-bracket reform in a neighborhood of y. The test for Pareto efficiency then makes use of the following insights:

1. There is no Pareto-improving one-bracket reform if and only if the function $y \mapsto \mathcal{R}(y)$ is bounded from below by 0 and from above by 1. These bounds

admit an interpretation as Laffer conditions which, respectively, indicate whether marginal tax rates are inefficiently high or inefficiently low.

2. There is no Pareto-improving two-bracket reform if and only if the function $y \mapsto \mathcal{R}(y)$ is non-increasing. A violation of this monotonicity condition implies that the tax system can be Pareto improved by an EITC-like reform – i.e. a two bracket reform with a phase-in and a phase-out range.

Thus, all that is needed to test for Pareto efficiency are sufficient-statistics formulas for the revenue implications of small tax reforms, each involving a modification of the marginal tax rate in a single bracket. The literature using perturbation methods in optimal taxation provides many examples of such sufficient-statistics formulas. Upon squaring our results with the formulas from that literature, one obtains a simple and complete test for Pareto efficiency. For concreteness, we present such sufficient-statistics formulas for a Mirrleesian model of income taxation with behavioral responses only at the intensive margin, and an extended model that also involves fixed costs of labor market participation.

When a tax system fails the Pareto test, this raises further questions. Is it possible to identify Pareto-improving direction for reform? Put differently, is it possible to identify "progress in tax policy"? Moreover, is it possible to measure how inefficient a given tax system is? Such a measure makes it possible to relate an inefficiency prevailing in one part of the tax system to another one prevailing in some other part of the tax system. For instance, a question that arises in the context of our application is whether the inefficiency in the tax system faced by childless singles was smaller than the inefficiency in the tax system for single parents.

We provide answers to these questions along the following lines: As we show, diagnosing whether an arbitrary tax perturbation is Pareto-improving is straightforward once an estimate of the reform's revenue implications is available. We also show how such an estimate can be obtained when the revenue implications of simple one-bracket reforms are known, or, equivalently, when the function $y \mapsto \mathcal{R}(y)$ is known. Thus, the information that is needed to test for Pareto efficiency can also be used to identify Pareto-improving directions for reform.

We also propose a money-metric measure of how inefficient a given tax system is. Specifically, we characterize the Pareto-improving direction for reform that yields the largest revenue gain above what is needed to ensure that no agent is made worse off. When this maximal revenue gain is small, the tax system is close to the Pareto frontier, otherwise it is not. Again, this measure can be computed when the function $y \mapsto \mathcal{R}(y)$ is known.

The introduction of the EITC. We look at the introduction of the EITC in the US through the lens of this framework. The introduction of the EITC in 1975 was a substantial policy change for many low-income households, see, e.g., Bastian (2020). It was meant as a response to excessively high marginal tax rates for families that depended on welfare. We use this setting as a testbed for our approach. Specifically, we derive the requisite sufficient statistics from a model with behavioral responses at the intensive and the extensive margin. Thereby, we obtain a test function $y \mapsto \mathcal{R}(y)$, which we then use to investigate whether or not the system prevailing at the time was Pareto-efficient, and whether reforms of the EITC in subsequent years went into a Pareto-improving direction.

We find that, prior to the introduction of the EITC, the function $y \mapsto \mathcal{R}(y)$ was increasing over certain income ranges, indicating the existence of a Pareto-improving two-bracket reform. The introduction of the EITC did not fully remove these inefficiencies, and left room for further Pareto improvements by means of two-bracket reforms. The first EITC expansion in 1979 then went into a Pareto-improving direction. These findings are shown to be robust with respect to alternative assumptions about the behavioral responses to taxation, in particular the extensive margin and intensive margin elasticities of labor supply. Thus, both the introduction and the subsequent expansion of the EITC can be rationalized through the lens of our framework.

The EITC and the theory of optimal taxation. Previous literature on the desirability of the EITC has used an optimal tax approach, thereby providing an answer to the following question: Are negative marginal taxes, or, equivalently, earnings subsidies for the "working poor" part of a tax policy that maximizes a utilitarian social welfare function? Providing a positive answer is not straightforward. The workhorse of the optimal tax literature, the Mirrlees (1971) model, stipulates non-negative marginal tax rates for all levels of income.[3] Thus, the EITC is a challenge for the theory of optimal taxation. In response to that challenge, Saez (2002) suggested the use of an extended version of the Mirrlees model that includes fixed costs of labor market participation and gives rise to behavioral responses both at the intensive and the extensive margin. With such a framework, there are conditions under which the EITC can be justified as being part of a welfare-maximizing policy, see Saez (2002) and Hansen (2021).[4]

Our analysis complements these findings by focusing on the tax and transfer system that prevailed when the EITC was introduced, and by taking a tax reform perspective. This relates our approach to a literature in public finance that emphasizes the analysis of reforms, i.e., of incremental changes of a given system, as opposed to an analysis of optimal tax systems.[5] With a tax reform perspective, we find that the EITC can be rationalized under weaker conditions than with an optimal tax perspective. First, we find that the introduction of the EITC was Pareto-improving, and not just utilitarian-welfare-improving. Second, when exploring alternative assumptions about intensive and extensive margin elasticities, we find that the EITC was Pareto-improving even without behavioral responses at the extensive margin. Thus, the introduction of the EITC was a good idea – even under the behavioral assumptions of the basic Mirrlees model.[6]

[3]Negative marginal tax rates can be rationalized in the basic version of the Mirrlees model only with a welfare function that has non-monotonic welfare weights, e.g., one that assigns higher weights to people with intermediate incomes than to people with low incomes; see Stiglitz (1982), Choné and Laroque (2010), or Brett and Weymark (2017).

[4]For a non-utilitarian assessment based on generalized weights, see Saez and Stantcheva (2016).

[5]See Piketty (1997), Saez (2001), Golosov et al. (2014), or Jacquet and Lehmann (2021). For earlier contributions to the analysis of tax reforms, see Feldstein (1976), Weymark (1981) and the review in Guesnerie (1995).

[6]Kleven (2019) suggests that previous estimates of extensive margin elasticities were too high.

Outline. The remainder is organized as follows. The next section discusses the related literature. Section 3 contains our theory of Pareto-improving tax reforms which yields a characterization of necessary and sufficient conditions for the Pareto efficiency of a tax system. Section 4 shows how can identify Pareto-improving reform directions. It also introduces a quantitative measure that makes it possible to compare different tax systems according to how inefficient they are. Section 5 contains the application of these conditions to the introduction of the EITC in the US. Formal proofs are relegated to the Appendix.

2 Related literature

We build on and extend the existing literature on Pareto-efficient non-linear taxation. Previous literature has generalized the notion of a Laffer bound to non-linear tax schedules, see Stiglitz (1982), Brito, Hamilton, Slutsky and Stiglitz (1990) and, more recently, Badel and Huggett (2017).[7] Bierbrauer, Boyer and Peichl (2021) show that there is not only an upper Pareto bound, but also a lower Pareto bound for marginal tax rates. This lower bound is relevant for an assessment of earnings subsidies: if the bound is violated, then a reduction of these subsidies is Pareto-improving. Werning (2007) and Lorenz and Sachs (2016) develop a test for the Pareto efficiency of a given status quo tax schedule that involves a differential equation that describes how marginal tax rates change along the income distribution.[8] Failures of Pareto efficiency are also identified by the literature on the

While this debate has a bearing on the desirability of the EITC from an optimal tax perspective, it is of no consequence for our conclusion that the introduction of the EITC was a reform that went into a Pareto-improving direction.

[7]There is a literature deriving the second-best Pareto frontier for a two-type Mirrlees model with contributions by Stantcheva (2014), Bierbrauer and Boyer (2014), and Bastani, Blomquist and Micheletto (2020). See, for reviews, Stiglitz (1987) and Boadway and Keen (2000).

[8]Werning (2007) considers a Mirrleesian framework where behavioral responses to taxation arise only at the intensive margin. Lorenz and Sachs (2016) consider, in addition, extensive margin responses. For related work, see also Blundell and Shephard (2012), Scheuer (2014), Koehne and Sachs (2019), or Hendren (2020).

inverse tax problem.[9] This literature identifies the welfare function that is maximized by an observed tax schedule. If this approach yields a welfare function with negative weights, this indicates that the tax policy under study is incompatible with the maximization of a Paretian social welfare function.[10]

All these contributions have in common that they focus on necessary conditions for Pareto efficiency. They do not provide sufficient conditions, i.e., there is no way of checking whether a given tax schedule satisfies *all* the conditions that are needed for Pareto efficiency. Instead, any one of these papers looks at a particular subset of these conditions. If the condition under consideration is violated, one can conclude that the given tax system can be reformed in a Pareto-improving way. If instead the condition is satisfied, one cannot conclude that the tax system is Pareto-efficient. The possibility of some other Pareto-improving reform remains.

In our approach, we consider an arbitrary number of brackets that can be distributed in an arbitrary way over the range of possible incomes. Allowing for a larger class of reforms than the previous literature enables us to show that, taken together, the conditions in Bierbrauer et al. (2021), on the one hand, and the conditions by Werning (2007) and Lorenz and Sachs (2016), on the other, imply Pareto efficiency.[11]

We moreover present sufficient-statistics formulas that can be used to check whether the necessary and sufficient conditions for the existence of Pareto-improving reform directions are satisfied. This relates our analysis to a broad literature employing sufficient statistics for policy evaluation, see Chetty (2009) and Kleven (2021) for reviews of this approach.

[9]See, e.g., Christiansen and Jansen (1978), Blundell, Brewer, Haan and Shephard (2009), Bourguignon and Spadaro (2012), Bargain, Dolls, Neumann, Peichl and Siegloch (2011), Jacobs, Jongen and Zoutman (2017), Lockwood and Weinzierl (2016), or Hendren (2020).

[10]In part E of the Supplementary Material, we explain in detail what the conditions for Pareto efficiency in Theorems 1 and 2 imply for the inverse tax problem.

[11]A qualification needs to be added: Werning (2007) and Lorenz and Sachs (2016) impose assumptions of differentiability that are not needed in our approach. We get back to the differential equations in Werning (2007) and Lorenz and Sachs (2016) after presenting our main results in Section 3.

Our analysis of the introduction of the EITC in the US draws on the literature that provides estimates of the behavioral responses involved.[12] There is a range of estimates and some of the conventional wisdom in the literature has recently been challenged.[13] Specifically, Bastian (2020) estimates the behavioral responses to the 1975 EITC introduction, the reform that we focus on as well. Bastian and Jones (2021) provide an econometric analysis of the extent to which EITC expansions since the 1990s were self-financing. Their analysis takes account of a wide range of potential fiscal externalities associated with an expansion of the EITC.

3 Pareto-improving tax reforms

In this section, we present results on Pareto-efficient income taxation and Pareto-improving tax reforms. These results are general in the sense that they are not tied to a specific setup, such as a Mirrleesian model or a model with fixed costs of labor market participation.

3.1 The model

We consider an economy with a continuum of individuals. Individuals value consumption c and generate earnings y. The generation of earnings comes with effort costs that depend on a vector of individual characteristics $\theta \in \Theta \subset \mathbb{R}^n$. The cross-section distribution of θ is assumed to be atomless and represented by a cumulative distribution function F. Preferences are represented by the utility function $u : \mathbb{R}^2_+ \times \Theta \to \mathbb{R}$. Thus, $u(c, y, \theta)$ is the utility that a type θ individual derives from a bundle (c, y). The function u is continuously differentiable

[12]Prominent references are Eissa and Liebman (1996), Meyer and Rosenbaum (2001), Moffitt (2003), Eissa and Hoynes (2004), and Blundell (2006). For surveys, see Hotz and Scholz (2003), Nichols and Rothstein (2015), and Hoynes (2019).

[13]For single mothers, early papers such as Meyer and Rosenbaum (2001) found large participation elasticities (sometimes above 1) along with only small responses at the intensive margin, while more recent papers such as Bastian (2020) and Kleven (2019) find smaller participation elasticities.

and increasing in the first argument, with partial derivative denoted by u_c. It is decreasing in the second argument y.

We assume that a single-crossing condition holds in one dimension of the type space, Θ_j: If type (θ_j, θ_{-j}) weakly prefers a bundle (c, y) to another bundle $(c', y') < (c, y)$, then type (θ'_j, θ_{-j}) with $\theta'_j > \theta_j$ strictly prefers (c, y) to (c', y'). This assumption implies that the individuals' earnings are increasing in θ_j.

Two special cases of this setup are of particular interest. First, a utility function that is quasilinear in consumption and has iso-elastic effort costs, in combination with a one-dimensional type space, i.e.,

$$u(c, y, \theta) = c - \frac{1}{1 + \frac{1}{\varepsilon}} \left(\frac{y}{\theta}\right)^{1+\frac{1}{\varepsilon}} ,$$

where $\theta \in \Theta \subset \mathbb{R}_+$. The analysis of Diamond (1998) is based on this framework. In this specification, the type θ is a measure of productive ability, often identified with an hourly wage. This case is of pedagogical interest. It is the simplest framework that we can use for purposes of illustration. Second, a model with multidimensional heterogeneity due to fixed and variable effort costs of productive effort, and behavioral responses both at the intensive and the extensive margin. The analysis in Section 4 is based on such a framework.

There is a status quo tax policy. It is represented by a parameter c_0 and a tax function T_0, which jointly define the budget set $C_0(y) = c_0 + y - T_0(y)$ that individuals face. The parameter c_0 is the intercept of this consumption schedule. It is the transfer to individuals with no earnings. Without loss of generality, we let $T_0(0) = 0$.[14] We assume that T_0 is continuous. Otherwise, it can be an arbitrary non-linear tax function, possibly with kinks. Before the reform, individuals solve

$$\max_{y \in \mathcal{Y}} \quad u(C_0(y), y, \theta) ,$$

where $\mathcal{Y} = [0, \bar{y}]$ is a set of feasible earnings level.

[14] In the literature, $T_0(y)$ is often referred to as the *participation tax*; see, e.g., Kleven (2014). This reflects that $T_0(y)$ is the additional tax payment of a person with earnings of y, relative to a person with no earnings. Alternatively, we could represent the status quo by a tax function $\tilde{T}_0$ so that $\tilde{T}_0(y) := -c_0 + T_0(y)$ with the implication that $\tilde{T}_0(0) = -c_0$. We find it more convenient to separate the transfer c_0 from the tax function.

A tax reform replaces T_0 by a new tax function T_1 so that $T_1 = T_0 + \tau\ h$. The scalar τ is a measure of the size of the tax reform and the function h gives the direction of the tax reform. Again, h is assumed to be a continuous function with $h(0) = 0$. For a given income y, the change in the tax burden due to the reform is therefore given by $T_1(y) - T_0(y) = \tau\ h(y)$. After the reform, individuals solve

$$\max_{y \in \mathcal{Y}} \quad u(C_1(y), y, \theta) \ , \tag{1}$$

where $C_1(y) = c_1 + y - T_0(y) - \tau\ h(y)$, and c_1 is the intercept after the reform. We denote the reform-induced changes in tax revenue by $R(\tau, h)$ and assume that it is absorbed by the intercept so that

$$c_1 = c_0 + R(\tau, h) \ .$$

Thus, any change in tax revenue is redistributed in a lump-sum fashion.

The change in tax revenue $R(\tau, h)$ is an endogenous object that depends on the behavioral responses to taxation. To see how it is determined, let $y^*(e, \tau, h, \theta)$ be the solution to (1), where

$$C_1(y) = c_0 + e + y - T_0(y) - \tau\ h(y) \ ,$$

and e is a source of income that is exogenous from the individual's perspective. Also, let $y_0(\theta) := y^*(0, 0, h, \theta)$ be a shorthand for income in the status quo.[15] Then, $R(\tau, h)$ solves

$$R(\tau, h) = \mathbf{E}\left[T_1(y^*(R(\tau, h), \tau, h, \theta)) - T_0(y_0(\theta))\right] \ , \tag{2}$$

where the operator $\mathbf{E}$ indicates that we compute a population average using the distribution F.[16]

We denote by $v(\tau, h, \theta)$ the indirect utility that a type θ individual realizes after a tax reform (τ, h). We can use the analysis of *"Envelope theorems for arbitrary*

[15] There may be types for whom the utility-maximization problem in (1) has multiple solutions. The function y^* is then taken to select one of them. How this selection is done is inconsequential for the analysis that follows.

[16] Brouwer's fixed point theorem can be used to establish the existence of a solution to this fixed point equation, in combination with conditions that ensure that $\mathbf{E}\left[T_1(y^*(e, \tau, h, \theta))\right]$ is continuous in e. This continuity is not immediate when the function y^* may exhibit jumps due

choice sets" in Milgrom and Segal (2002) to describe how individuals are affected by marginal changes of the reform intensity τ. Specifically, fix some type θ. Then, by Corollary 4 in Milgrom and Segal (2002),

$$\frac{d}{d\tau}\, v(\tau, h, \theta) = u_c(\cdot, \theta)\ \left[R_\tau(\tau, h) - h\left(y^*(\cdot)\right)\right]\ , \tag{3}$$

where the marginal consumption utility of type θ, $u_c(\cdot, \theta)$, is evaluated at point $(C_1(y^*(\cdot)), y^*(\cdot))$, and $R_\tau(\tau, h)$ is the marginal effect of an increase in the reform intensity τ on tax revenue.[17] More formally, it is the Gateaux differential of tax revenue in direction h.[18] The envelope theorem covers cases in which the marginal tax rates (either in the status quo or after the reform) exhibit discontinuous jumps. It also applies when there are fixed costs of labor market participation, so that the utility function is, at $y = 0$, not continuous in y.

Equation (3) makes it possible to decompose the set of taxpayers into winners and losers of the tax reform. For concreteness, fix a reform direction h and suppose that, starting from the status quo policy, a small reform step has a positive impact on tax revenue, $R_\tau(0, h) > 0$. A taxpayer with type θ benefits from the reform if and only if this revenue gain outweighs the additional tax payment $h(y_0(\theta))$.

to extensive margin responses or discontinuities in marginal tax rates. With a single crossing condition on preferences and an atomless type distribution, however, such jumps can be shown to wash out in the aggregate and therefore do not upset the continuity of $\mathbf{E}\left[T_1(y^*(e, \tau, h, \theta))\right]$ in e.

[17] For a type θ such that the utility-maximization problem in (1) has multiple solutions, the right-hand derivative of v is relevant for increases of τ and the left-hand derivative is relevant for decreases of τ.

[18] Our notation for Gateaux differentials is inspired by the one for partial derivatives. Conventions in mathematics are different. To make this explicit, let tax revenue $\mathcal{R}$ be a real-valued functional of the tax function T. Then, the Gateaux differential of tax revenue in direction h is formally defined as

$$\partial\mathcal{R}(T, h) := \lim_{\tau \to 0} \frac{\mathcal{R}(T + \tau\, h) - \mathcal{R}(T)}{\tau}\ ,$$

where the left-hand side is the "typical" notation in the literature. Our notation can now be more formally introduced as $R_\tau(0, h) := \frac{d}{d\tau}\mathcal{R}(T_0 + \tau h)|_{\tau=0} = \partial\mathcal{R}(T_0, h)$. In Appendix A, we lay down further assumptions which guarantee the linearity of the Gateaux differential in the direction h, a property that is used in the proofs of Theorems 1 and 2.

Hence, given reform direction h, a small increase of τ is Pareto-improving if

$$R_\tau(0, h) - \max_{y \in y_0(\Theta)} h(y) > 0 \ , \tag{4}$$

where $y_0(\Theta)$ is the image of function y_0.[19]

We say that there is no Pareto-improving direction in a class of reforms H if, for all functions $h \in H$,

$$R_\tau(\tau, h) - \max_{y \in y^*(\Theta)} h(y) < 0 \ . \tag{5}$$

The set H will be expanded as we go along. We first analyze classes of reforms with one or two income brackets in which marginal tax rates are changed. We then extend the results to tax reforms with finitely many brackets and, finally, cover the entire set of continuous reform directions h.

Single-bracket reforms. A single-bracket reform is a pair (τ, h^s), where the function h^s is such that

$$h^s(y) = \begin{cases} 0, & \text{if} \quad y \leq \hat{y} \ , \\ y - \hat{y}, & \text{if} \quad y \in (\hat{y}, \hat{y} + \ell) \ , \\ \ell, & \text{if} \quad y \geq \hat{y} + \ell \ . \end{cases}$$

for some threshold value of income $\hat{y}$, see Figure 1. Thus, a single-bracket reform is characterized by a triplet $(\tau, \ell, \hat{y})$, where $\hat{y}$ is the income level at which the bracket starts, ℓ is the length of the bracket and τ is the amount by which the marginal tax rate changes for incomes in the bracket.

After a one-bracket reform, the new tax schedule is given by

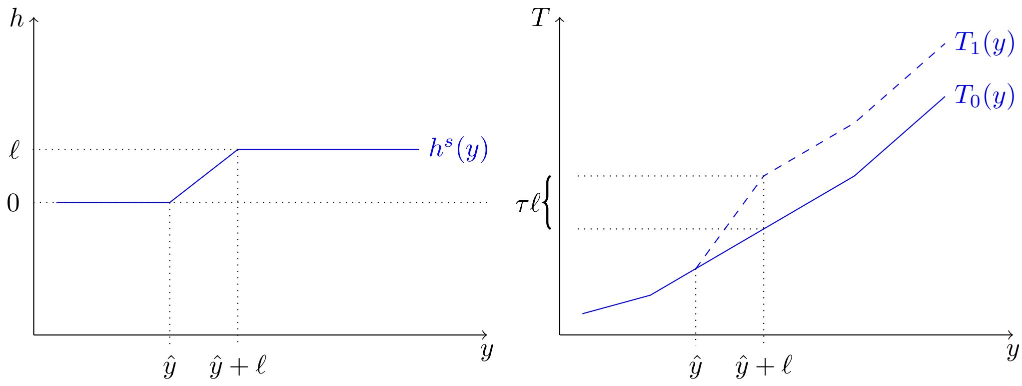
$$T_1(y) = T_0(y) + \tau h^s(y) = \begin{cases} T_0(y), & \text{if} \quad y \leq \hat{y} \ , \\ T_0(y) + \tau\,(y - \hat{y}) \ , & \text{if} \quad y \in (\hat{y}, \hat{y} + \ell) \ , \\ T_0(y) + \tau \ell, & \text{if} \quad y \geq \hat{y} + \ell \ . \end{cases}$$

Hence, the reform increases tax liabilities for all earnings above $\hat{y}$, with a maximum increase of $\tau\ell$. The marginal tax rate changes by τ for earnings in the bracket

[19] A small reduction of τ for direction h is Pareto-improving if $R_\tau(0, h) - \min_{y \in y_0(\Theta)} h(y) < 0$, or, equivalently, if a small increase of τ for direction $-h$ is Pareto-improving.

$(\hat{y}, \hat{y}+\ell)$. It does not change for incomes above or below this bracket. Formally, the new schedule of marginal tax rates equals

$$T_1'(y) = T_0'(y) + \tau h^{s\prime}(y) = \begin{cases} T_0'(y), & \text{if} \quad y \leq \hat{y} \ , \\ T_0'(y) + \tau, & \text{if} \quad y \in (\hat{y}, \hat{y}+\ell) \ , \\ T_0'(y), & \text{if} \quad y \geq \hat{y}+\ell \ . \end{cases}$$



(a) Reform direction h^s.

(b) Tax functions T_0 and T_1.

Figure 1: Income tax reforms with one bracket.

We will trace the welfare implications of multi-bracket reforms back to the properties of single-bracket reforms. It will prove convenient to have separate notation for the revenue implications of single-bracket reforms. For such reforms, we write $R^s(\tau, \ell, \hat{y})$ rather than $R(\tau, h^s)$. We write R^s_τ for the derivative of this function with respect to the first argument, and $R^s_{\tau\ell}$ for the cross-derivative with respect to the first and the second argument. It follows from first-order Taylor approximations that, for τ and ℓ close to zero,

$$\tau \ \ell \ R^s_{\tau\ell}(0, 0, y)$$

is a good approximation of $R^s(\tau, \ell, y)$, i.e., of the reform's revenue implications. Thus, the cross-derivative $R^s_{\tau\ell}$ can be interpreted as a measure of how much revenue can be raised by a small single-bracket reform. The function $y \mapsto R^s_{\tau\ell}(0, 0, y)$ is a recurrent theme in what follows. For a more concise notation, we will henceforth write $\mathcal{R}(y)$ rather than $R^s_{\tau\ell}(0, 0, y)$ and frequently refer to the function $y \mapsto \mathcal{R}(y)$.

Two-bracket reforms. A two-bracket reform combines two single-bracket reforms. Formally, it is a pair (τ, h_2), where the subscript of h_2 signifies a reform involving two brackets. The function h_2 is defined by

$$h_2(y) := \tau_1 \, h_1^s(y) + \tau_2 \, h_2^s(y) \; , \tag{6}$$

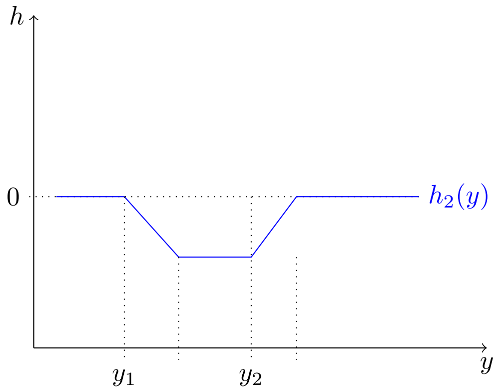
for

$$h_1^s(y) = \begin{cases} 0, & \text{if} \quad y \leq y_1 \; , \\ y - y_1, & \text{if} \quad y \in (y_1, y_1 + \ell \, \ell_1) \; , \\ \ell \, \ell_1, & \text{if} \quad y \geq y_1 + \ell \, \ell_1 \; . \end{cases}$$

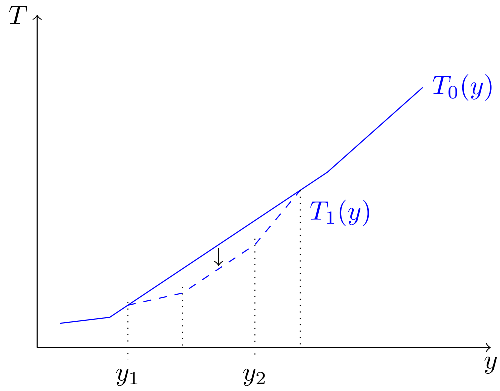
and

$$h_2^s(y) = \begin{cases} 0, & \text{if} \quad y \leq y_2 \; , \\ y - y_2, & \text{if} \quad y \in (y_2, y_2 + \ell \, \ell_2) \; , \\ \ell \, \ell_2, & \text{if} \quad y \geq y_2 + \ell \, \ell_2 \; . \end{cases}$$

Thus, a two-bracket reform links two single-bracket reforms in a particular way: marginal tax rates change by $\tau \; \tau_1$ for incomes in the first bracket and by $\tau \; \tau_2$ for incomes in the second bracket. The first bracket has a length of $\ell \, \ell_1$, and the second bracket has a length of $\ell \, \ell_2$.



(a) Reform direction h_2.



(b) Tax functions T_0 and T_1.

Figure 2: A two-bracket tax cut.

The new tax schedule satisfies

$$T_1(y) = T_0(y) + \tau \; h_2(y) \; ,$$

and the new schedule of marginal tax rates equals

$$T_1'(y) = T_0'(y) + \tau\, h_2'(y) \ ,$$

where

$$h_2'(y) = \begin{cases} \tau_1, & \text{for} \quad y \in (y_1, y_1 + \ell\, \ell_1) \ , \\ \tau_2, & \text{for} \quad y \in (y_2, y_2 + \ell\, \ell_2) \ , \\ 0, & \text{for} \quad y \leq y_1 \ , y \in [y_1 + \ell\, \ell_1, y_2] \ , y \geq y_2 + \ell\, \ell_2 \ . \end{cases}$$

In what follows, two-bracket reforms with $\tau_1 < 0$, $\tau_2 > 0$ and $\tau_1 \ell_1 + \tau_2 \ell_2 = 0$ are of particular interest. We refer to these reforms as *two-bracket tax cuts.* This choice of terminology reflects that these reforms do not increase anyone's tax burden and that all people with an income between the endpoints of the two brackets get a tax cut. Moreover, they involve a phase-in range where marginal taxes are reduced, and a subsequent phase-out range where marginal taxes are increased, see Figure 2.

Our construction of two-bracket reforms facilitates an analysis of the limit case $\tau \to 0$ and $\ell \to 0$, see Figure 3 for the case of a small two-bracket tax cut. As τ goes to zero, the ratio of the marginal tax rate changes is kept constant at $\frac{\tau_1}{\tau_2}$. Analogously, both brackets shrink when ℓ is send to zero, while the ratio of their lengths is kept constant at $\frac{\ell_1}{\ell_2}$.

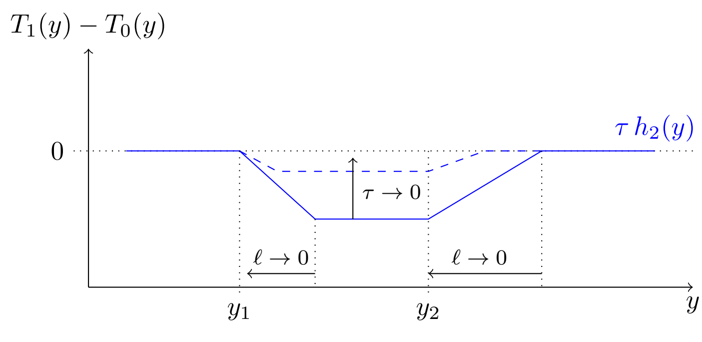


Figure 3: A small two-bracket tax cut.

Reforms with finitely many brackets. We extend the construction of two-bracket reforms to reforms with a finite number of brackets in the natural way: A reform (τ, h_m) with m brackets is given by a collection of m single-bracket reforms such that $h_m(y) := \sum_{k=1}^{m} \tau_k \; h_k^s(y)$, where

$$h_k^s(y) = \begin{cases} 0, & \text{if} \quad y \leq y_k \; , \\ y - y_k, & \text{if} \quad y \in (y_k, y_k + \ell \; \ell_k) \; , \\ \ell \; \ell_k, & \text{if} \quad y \geq y_k + \ell \; \ell_k \; . \end{cases}$$

The reform induced tax schedule is then given by $T_1 = T_0 + \tau \; h_m$.

3.2 The main results

Under what conditions is it possible to make everyone better off by increasing or lowering the marginal tax rates in a finite number of income brackets? Theorems 1 and 2 provide answers to this question.

Theorem 1 *If T_0 is a Pareto-efficient tax system, then the function $y \mapsto \mathcal{R}(y)$ is bounded from below by 0, bounded from above by 1, and non-increasing.*

Theorem 1 states necessary conditions for the Pareto efficiency of a tax system. The first condition is that $\mathcal{R}(y) \geq 0$ for all y. Hence, a one-bracket reform involving an increase of marginal tax rates must not lead to a loss of tax revenue. If the condition was violated, it would be possible to raise revenue by means of a tax cut, and such a reform would be Pareto-improving. The logic is familiar from analyses of the Laffer curve. The second condition is that $\mathcal{R}(y) \leq 1$ for all y. It is a mirror image of the first condition. If it was violated, it would be possible to raise so much revenue by increasing marginal tax rates that even those who suffer most from the tax increase would be compensated. If T_0 is a Pareto-efficient tax system, there must be no scope for such a Pareto improvement.

The following Proposition clarifies what reform options exist when the function $y \mapsto \mathcal{R}(y)$ is increasing over some range.

Proposition 1 *If there are two income levels y_1 and $y_2 > y_1$ such that $\mathcal{R}(y_2) > \mathcal{R}(y_1)$, there exists a Pareto-improving two-bracket tax cut with $\tau > 0$ and $\ell > 0$.*

In light of Proposition 1, Theorem 1 provides a characterization of two necessary conditions for Pareto efficiency: first, there must not be a Pareto improvement in the class of one-bracket reforms. Second, there must not be a Pareto improvement in the class of two-bracket tax cuts. As we show formally in the proof of Proposition 1, if $y \mapsto \mathcal{R}(y)$ is increasing, a two-bracket tax cut between incomes y_1 and y_2 is self-financing: the revenue loss due to a reduction of marginal tax rates in the first bracket is more than offset by the revenue gain from the increase of marginal tax rates in the second bracket. Thus, the condition that $y \mapsto \mathcal{R}(y)$ must be non-increasing is an analogue to the condition $\mathcal{R}(y) \geq 0$ for all y. The latter rules out self-financing tax cuts for one-bracket reforms. The former does so for two-bracket reforms.

Theorem 1 and Proposition 1 show that there may exist Pareto-improving two-bracket reforms, even when no Pareto-improving one-bracket reform can be found. Given this finding, one might conjecture that there is no hope to obtain a concise characterization of Pareto-efficient tax systems: even if one found a condition ruling out Pareto-improving two-bracket reforms, there would still be the possibility of a Pareto-improving three-bracket reform. If one had eliminated those, one would still have to deal with four-bracket reforms, and so on. Theorem 2 shows that this is not the case: ruling out Pareto-improving one- and two-bracket reforms is sufficient for Pareto efficiency.

Theorem 2 *If the function $y \mapsto \mathcal{R}(y)$ is bounded from below by 0, bounded from above by 1, and non-increasing then there is no Pareto-improving direction in the class of reforms with finitely many brackets.*

According to Theorem 2, ruling out Pareto-improving one- and two-bracket reforms guarantees Pareto efficiency. The Theorem implies, in particular, that any tax system that can be Pareto-improved by a tax reform that affects three or more brackets, can also be Pareto-improved by a tax reform that affects at most two brackets.

Why is Theorem 2 true? The class of m-bracket reforms gives rise to many possible combinations of rate increases and rate cuts. There could be increases in all odd brackets and decreases in all even brackets; there could be increases for high incomes and decreases for low incomes, etc. The difficulty in the proof is to show that none of these combinations can be Pareto-improving under the conditions of Theorem 2. Here we give an intuition that covers some of these cases.

Assume first that marginal tax rates go up in the first bracket. With $\mathcal{R}$ throughout below 1, this yields a revenue gain that is, however, not large enough to compensate people with incomes above the first bracket who now face an increased tax burden. Thus, additional revenue is needed and this requires to increase marginal taxes in some other bracket k higher up in the income distribution. With $\mathcal{R}(y) < 1$, if we raise enough revenue in bracket k so as to compensate the people below, this will only add to making people with incomes above bracket k worse off, and so on if we add further brackets in which taxes go up. Moreover, there is no way to overcome this by mixing in brackets with tax cuts. When $\mathcal{R}(y) > 0$, for all y, this only aggravates the difficulty of drumming up enough revenue for the compensation of those who face higher taxes. Thus, there is no Pareto improvement with a rate increase in bracket 1.

How about having instead a first bracket with a tax cut? People with incomes above this bracket are then made better off; but, with $\mathcal{R}$ positive throughout, this creates a revenue loss that is harmful for anyone else. Adding further brackets where tax rates are lowered leads to an even more substantial revenue loss. We can now try to offset this effect with brackets in which tax rates go up. Assume that we increase marginal taxes in the second bracket just enough to make sure that agents with incomes above this bracket face neither a tax cut not a tax increase, $\tau_2 \ell_2 = -\tau_1 \ell_1$ (hence, we employ a two-bracket tax cut). With $y \mapsto \mathcal{R}(y)$ decreasing, the revenue gain due the second bracket is not large enough to make up for the revenue loss from the first bracket. Hence, we have to raise marginal taxes in the second bracket – or in further brackets – even more. But then, the reform raises the overall tax burden for people further up in the income distribution. This neither is Pareto-improving. Hence, irrespectively of whether we have a tax increase or a tax cut in the first bracket, those who are hit hardest by an increase of their tax liability cannot be made better off.

Continuous reform directions. So far, our results were restricted to the class of tax reforms with finitely many brackets, thereby excluding, e.g., continuously differentiable reform directions. The following corollary extends Theorem 2 to cover the entire class of continuous reform directions. It exploits that any continuous function $h : \mathcal{Y} \to \mathbb{R}$ can be approximated arbitrarily well by an m-bracket reform with m sufficiently large.

Corollary 1 *If $y \mapsto \mathcal{R}(y)$ is bounded from below by 0, bounded from above by 1, and non-increasing, there is no Pareto-improving direction in the class of continuous functions $h : \mathcal{Y} \to \mathbb{R}$.*

According to Corollary 1, any tax system that can be Pareto-improved by any continuous reform, can also be Pareto-improved by a tax reform that affects at most two brackets.

An alternative characterization. As we discuss in part D of the Online Appendix, when earnings are bounded away from zero and bounded from above, it is possible to obtain a more parsimonious characterization of sufficient conditions for Pareto efficiency: The monotonicity condition on $y \mapsto \mathcal{R}(y)$ is then sufficient for Pareto efficiency.[20] This said, for our application of interest, the introduction of the EITC, incentives for labor market participation play a key role. In this context, an assumption that everybody has strictly positive earnings would not be appropriate.

[20]We are grateful to an anonymous referee for pointing to this possibility.

3.3 Discussion

Sufficient statistics for $\mathcal{R}$. All previous results are expressed using the function $y \mapsto \mathcal{R}(y)$. Thus, given an estimate of this function, our approach allows to check whether a tax system of interest is Pareto-efficient. Different models of taxation give rise to different versions of the function $y \mapsto \mathcal{R}(y)$. The concrete specification will depend on the application of interest and on a choice of what model to use for this application. We illustrate this with two examples. The formulas that follow are derived in part F of the Supplementary Material.

First, consider the model of Diamond (1998) with $u(c, y, \theta) = c - \frac{1}{1+\frac{1}{\epsilon}} \left(\frac{y}{\theta}\right)^{1+\frac{1}{\epsilon}}$, where $\theta \in \Theta \subset \mathbb{R}_+$ is a measure of productivity and the parameter ϵ pins down the labor supply elasticity at the intensive margin. For this model, the revenue implications of a small one-bracket reform at income y are given by

$$\mathcal{R}(y) = 1 - F_y(y) - \varepsilon_0(y)\; y\; f_y(y)\; \frac{T_0'(y)}{1 - T_0'(y)} \;\;, \tag{7}$$

where F_y ist the *cdf* of the earnings distribution, and $\varepsilon_0 : y \mapsto \varepsilon_0(y)$ is a function which gives, for each level of y, the intensive-margin elasticity of earnings with respect to the retention rate $1 - T_0'(y)$.

Second, the literature on the desirability of earnings subsidies for the "working poor" suggests the use of a framework with taxpayers who differ both in the variable costs of productive effort and in the fixed costs of labor market participation.[21] We present a general framework and derive a sufficient-statistics formula characterizing $y \mapsto \mathcal{R}(y)$ in the Supplementary Material (see Proposition 4 in part F).[22] For ease of exposition, we focus here on the case of quasi-linear preferences and iso-elastic effort costs. This is also the specification that we will use in our

[21] A similar framework is also used in the literature on optimal pension and retirement policies, see, e.g., Golosov, Shourideh, Troshkin and Tsyvinski (2013), Michau (2014), and Shourideh and Troshkin (2017).

[22] This derivation is of stand-alone-interest in that it is based on a general specification of preferences, allowing for income effects, monetary or psychic fixed costs of labor market participation and complementarities between consumption and leisure.

benchmark analysis of the EITC in the subsequent section. Hence, suppose that

$$u(c, y, \omega, \gamma) = c - \frac{1}{1+\frac{1}{\epsilon}} \left(\frac{y}{\omega}\right)^{1+\frac{1}{\epsilon}} - \gamma \, \mathbb{1}_{y>0} \ ,$$

where ω and γ are, respectively, interpreted as a taxpayer's variable and fixed cost type. Thus, an individual's type θ is now taken to be a pair $\theta = (\omega, \gamma)$ and $\Theta = \Omega \times \Gamma$. Then,

$$\mathcal{R}(y) = 1 - F_y(y) - \varepsilon_0(y) \; y \; f_y(y) \; \frac{T_0'(y)}{1 - T_0'(y)} - \int_y^\infty f_y(y') \, \pi_0(y') \, \frac{T_0(y')}{y' - T_0(y')} \, dy' \ , \quad (8)$$

where $\pi_0(y)$ is an extensive-margin (participation) elasticity. It measures the percentage of individuals with an income of y who leave the labor market when their after-tax income $y - T_0(y)$ is decreased by one percent. This formula has first been derived by Jacquet et al. (2013) and also appears in Lorenz and Sachs (2016).

A further comment on related literature. Both Werning (2007) and Lorenz and Sachs (2016) present tests for Pareto efficiency that involve differential inequalities. Their approach can be illustrated with the setup of Diamond (1998). Recall that, in this case, $\mathcal{R}(y)$ is given by equation (7). If the expressions on the right hand side of (7) are taken to be differentiable in y, then the condition that $y \mapsto \mathcal{R}(y)$ must be non-increasing can be formulated as a differential equation that involves the derivative of the ratio $\frac{T_0'(y)}{1-T_0'(y)}$ and the derivative of the inverse hazard rate $\frac{f_y(y)}{1-F_y(y)}$. Both Werning (2007) and Lorenz and Sachs (2016) present such equations. Their findings are implied by our result in Theorem 1 that the monotonicity of the function $y \mapsto \mathcal{R}(y)$ is necessary for Pareto efficiency.

Tagging. Our analysis can be extended to allow for tagging.[23] Suppose that the population can be divided into separate groups and that it is publicly observable to which group a person belongs. The tax and transfer system may then treat individuals who belong to different groups differently. For instance, transfers and earnings subsidies for lone mothers may be larger than those for childless individuals. The above analysis of Pareto-efficient taxation can then be applied separately

[23] The seminal reference is Akerlof (1978). For a review, see Piketty and Saez (2013).

for each group. This implies, in particular, that revenue changes due to a tax reform that affects one group are rebated lump sum in this group.[24]

4 Evaluating tax reforms

Below, we apply the necessary and sufficient conditions for the Pareto efficiency of a tax system to the introduction of the EITC in the 1970s. We will find that there were indeed inefficiencies in the tax-and-transfer system prevailing at the time. Such a finding raises further questions: Did the 1975 EITC reform have a Pareto-improving direction? Relatedly, did the reform make the pre-existing inefficiency smaller? These questions require tools for an evaluation of tax reforms. We introduce them in this section. Proposition 2 below clarifies how one can check whether a tax reform has a Pareto-improving direction. We, moreover, introduce a money-metric measure of how inefficient a tax system is. This measure can be used to substantiate a statement such as "the inefficiency in the tax-and-transfer system for childless singles is small in comparison to the inefficiency prevailing in the one for single parents."

Pareto-improving directions. The following Proposition clarifies how one can check whether a specific tax reform has a Pareto-improving direction. It can also be used to evaluate whether reforms observed in the past had a Pareto-improving direction.

Proposition 2 *A reform in direction h is Pareto-improving if*

$$R_\tau(0,h) = \int_{\mathcal{Y}} h'(y)\ \mathcal{R}(y)\ dy \quad > \quad \max_{y \in y_0(\Theta)} h(y)\ . \tag{9}$$

Once a reform direction h is specified, inequality (9) can be used to determine whether it is Pareto-improving. The equation on the left-hand side shows that,

[24] This is without loss of generality. Redistributing the revenue gains from a self-financing tax reform among various groups can only make it more difficult to realize a Pareto improvement. There are then less resources that can be used to compensate those adversely affected by the tax reform.

with the function $y \mapsto \mathcal{R}(y)$ at hand, one obtains an estimate of the reform's impact on tax revenue in a straightforward way: it is a weighted average of the revenue implications of one-bracket reforms where the weights are given by the reform-induced changes of marginal tax rates, $y \mapsto h'(y)$. A reform with direction h and step size τ changes the marginal tax at income y by $\tau h'(y)$; hence, $h'(y)$ is the change in marginal tax rates per unit change of τ. According to (9), if the overall revenue gain is large enough to compensate even those agents who face the largest tax increase, then direction h is Pareto-improving.

Measuring the size of inefficiencies. When a tax system (T_0, c_0) is not Pareto-efficient, there is a set of Pareto-improving reform directions. Is there a way to judge whether one these reform directions is *better* than another? One possibility is to order reform directions according to $R_\tau(0,h) - \max_{y \in y(\Theta)} h(y)$, the tax revenue gain in excess of what is needed to compensate the agents facing the largest tax increase.[25] With this measure, the *best* reform direction h maximizes

$$\int_0^{\bar{y}} h'(y)\mathcal{R}(y)dy - \max_{y \in y_0(\Theta)} \int_0^y h'(z)dz$$

over the set of continuous functions $h : \mathcal{Y} \to \mathbb{R}$. This problem is linear in the function h' which gives the reform-induced change in marginal tax rates. So, to ensure the existence of a solution we also impose the constraint that the changes in marginal tax rates must be bounded, $|h'(y)| \leq a$ for each $y \in \mathcal{Y}$. Henceforth, h_R denotes the solution to this problem and we refer to it as the *revenue-maximizing tax cut.*

Proposition 3 in Appendix B characterizes the revenue-maximizing tax cut for a tax system that is inefficient because the revenue function $y \mapsto \mathcal{R}(y)$ is increasing over some income range.[26] As we show, the best reform direction h_R is then a two-bracket tax cut that affects marginal tax rates over an income range $[y_s, y_t]$. Specifically, it reduces marginal taxes in a phase-in range, going from y_s

[25] Tax revenue is a metric that is used in consumer choice theory when evaluating policy or price changes using compensating or equivalent variations. It is also used in Werning (2007)'s approach to Pareto-efficient taxation.

[26] This is the relevant scenario in our application to the 1974 US tax-and-transfer system below.

to $\frac{1}{2}(y_s + y_t)$, and increases marginal taxes in a phase-out range that begins at $\frac{1}{2}(y_s + y_t)$ and ends at y_t.[27] The impact on tax revenue can be written as

$$a\ I(T_0, c_0) = a\ \left[\int_{\frac{1}{2}(y_s+y_t)}^{y_t} \mathcal{R}(y)dy - \int_{y_s}^{\frac{1}{2}(y_s+y_t)} \mathcal{R}(y)dy \right]\ . \tag{10}$$

It equals the difference between the revenue gain in the phase-out range and the revenue loss in the phase-in range. Note that the term denoted by $I(T_0, c_0)$ does not depend on the parameter a that bounds the change in marginal tax rates. Hence, $I(T_0, c_0)$ admits an interpretation as a scale-invariant measure of how inefficient the tax system is. We can therefore use it to rank different tax systems according to how inefficient they are: Tax system (T_A, c_A) is more inefficient than tax system (T_B, c_B) if $I(T_A, c_A) > I(T_B, c_B)$. Thus, the more inefficient a tax system, the larger are the potential revenue gains from reforming it.[28]

Formally, $I(T_0, c_0)$ gives the slope of net revenue in direction h_R. For later reference, we note that, as an implication, a Taylor approximation can be used to obtain an estimate of the revenue implications of a reform (τ, h_R) with τ strictly positive, but close to zero:

$$R(\tau, h_R) - \max_{y \in y_0(\Theta)} \tau\ h_R(y) \simeq \tau\ a\ I\left(T_0, c_0\right)\ . \tag{11}$$

We apply this formula in the subsequent section where we discuss how close the actual reforms of the EITC came to the benchmark h_R. We then look at reforms that change marginal taxes rates by at most one percentage point. This requires to choose the parameters τ and a such that $\tau\ a = 0.01$.

5 Application: The introduction of the EITC

We now relate our insights on Pareto-improving tax reforms to the 1975 introduction of the EITC and its subsequent expansion. After describing the EITC

[27]The income levels y_s and y_t are optimally chosen. As we show in Appendix B, optimality requires that $\mathcal{R}(y_s) = \mathcal{R}(\frac{y_s+y_t}{2}) = \mathcal{R}(y_t)$.

[28]The revenue gain $I(T_0, c_0)$ can also be interpreted as a measure of distance from the Pareto frontier: it is zero for any Pareto-efficient tax system, and positive for any inefficient tax system.

1975 reform, we first use Theorem 1 in combination with the sufficient-statistics formula in (8) to show that the US tax-transfer system was not Pareto-efficient prior to the introduction of the EITC in 1974. We then apply Proposition 2 to check whether the direction of the 1975 EITC reform was Pareto-improving, and compare it to the revenue-maximizing tax cut for single parents at the time. We also show that, based on our inefficiency measure, the 1974 tax system for single parents was more inefficient than the one for childless singles.

5.1 Background on the EITC

The introduction of the EITC in 1975 was a response to a "poverty trap". In the 1960s, new welfare programs had been introduced as part of President Johnson's "war on poverty". On the one hand, the new programs provided more generous benefits to families with low incomes. On the other hand, these benefits were phased out in a way that implied high effective marginal tax rates for many low-income families, exceeding 70% in many cases (see Figure 4 below). In the following decade, the share of welfare recipients increased substantially. By the early 1970s, finding ways out of the "poverty trap" by an increase of work incentives was considered a pressing concern.[29]

The US Congress enacted the EITC as a temporary policy for the year 1975.[30] As described in Bastian (2020), this was a substantial policy change that affected a large share of the population.[31] It was set up as a refundable tax credit that was phased in at a marginal rate of 10% for taxpayers with less than 4,000 USD annual income, giving a maximum credit of 400 USD. The credit was then phased out at a marginal rate of 10% for incomes between 4,000 and 8,000 USD. Taxpayers with incomes above 8,000 USD were not eligible. The program was initially restricted to

[29]Detailed reviews of the debates at the time can be found in Ventry (2000), Moffitt (2003), or Nichols and Rothstein (2015).

[30]While the program was initially introduced under the name *Earned Income Credit*, it was soon relabeled to its current name *Earned Income Tax Credit*.

[31]According to CPS data, about 50% of single parents and 43% of childless singles had earned incomes in the EITC range (i.e., strictly positive and below $8,000$ USD).

working taxpayers with dependent children. Later, the EITC became a permanent policy. Over the following decades, there were several expansions.[32]

5.2 Calibration

We focus on two subgroups of the population, single parents and childless singles. In 1975, the EITC was introduced for the former, but not for the latter. Our analysis below will rationalize this policy choice: we will show that there was clearly scope for a Pareto-improving reform of the tax and transfer system for single parents, whereas no equally strong case can be made for childless singles. Our benchmark analysis is based on the formula for $y \mapsto \mathcal{R}(y)$ in equation (8).[33]

We use data on the most important elements of the US tax-and-transfer system for the tax year 1974 and later (see Table 1 in part C of the Supplementary Material for details). Specifically, we take account of the federal income tax and the two largest welfare programs, Aid for Families with Dependent Children (AFDC) and Supplementary Nutrition Assistance Programs (SNAP, also known as Food Stamps). The details of AFDC varied across states, so that a unified treatment for the US at large is not possible. We therefore focus on California, the state with the largest population both in the 1970s and today. Moreover, taxes and welfare transfers differed with respect to the number of children. In the following, we focus on the subgroup of single parents with two children.[34]

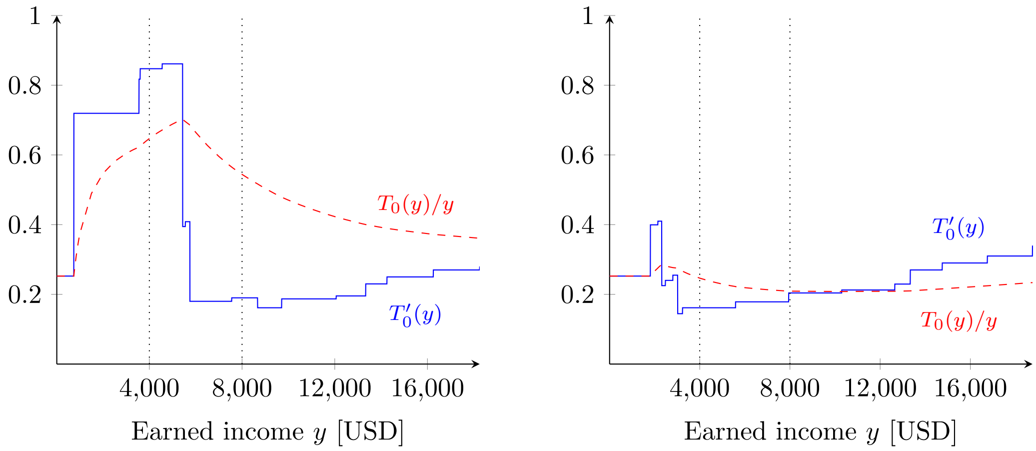
Figure 4 shows effective marginal tax rates, $y \mapsto T_0'(y)$, and participation tax rates, $y \mapsto \frac{T_0(y)}{y}$, for single parents (left panel) and for childless singles (right

[32]For example, more generous credits for parents with two or more children in 1991, and for parents with three or more children in 2009. In 1994, US authorities also introduced a more modest EITC for childless workers. In 2002, the eligibility thresholds were expanded for married taxpayers. See Hoynes (2019) for a review.

[33]This formula is based on a specification of preferences so that there are no income effects. We allow for income effects when we explore the robustness of our findings.

[34]In our data, the median number of children in single-parent households was two, while the arithmetic mean was 2.2. As we show in part C of the Supplementary Material, an analysis for the subgroups of parents with less or more children yields similar conclusions.

panel) before the reform in 1974.[35] At low incomes, both marginal tax rates and participation tax rates were much higher for single parents than for childless singles. The reason is that, for single parents, the phasing-out of AFDC and SNAP transfers implied an income range with exceptionally high marginal tax rates well above 70% and participation tax rates above 60%. This was not the case for childless singles. The dotted vertical lines in both panels of Figure 4 indicate the income range that was affected by the introduction of the EITC in 1975: It reduced marginal taxes in the phase-in range between 0 and 4,000 USD (first dotted line) and raised marginal taxes in the phase-out range between 4,000 and 8,000 USD (second dotted line).



(a) Single parents. (b) Childless singles.

Figure 4: US tax-transfer schedules 1974, single parents and childless singles.

Notes: Figure 4 shows the 1974 effective marginal tax $T_0'(y)$ (blue lines) and participation tax rate $T_0(y)/y$ (red lines) for single parents (left panel) and for childless singles (right panel) as functions of earned income in 1974 USD. The introduction of the EITC in 1975 decreased marginal taxes between 0 and 4,000 USD (first dotted vertical line) and increased them between 4,000 and 8,000 USD (second dotted vertical line).

Source: Authors' calculations (see part C of the Supplementary Material for details).

We estimate the 1974 income distributions in both subgroups based on data

[35]Recall that we define $T_0(y)$ to capture the participation tax at income y, i.e., the tax payment at income y relative to the tax payment at zero income. In the literature, the ratio $T_0(y)/y$ is commonly referred to as the participation tax rate, see, e.g., Kleven (2014).

from the March 1975 Current Population Survey (CPS). For this purpose, we consider the sample of non-married individuals aged 25 to 60 who do neither co-habit with an unmarried spouse nor with another adult family member. We partition this sample into childless singles and single parents.[36] In line with the EITC rules, we consider as earned income the sum of wage income and self-employment income. Single parents with strictly positive earned incomes below $8,000$ USD were eligible for the EITC.[37] For our benchmark analysis, we estimate the income distributions for both groups using a non-parametric kernel density estimation.[38]

We draw on a rich literature providing estimates of labor supply elasticities. Our benchmark analysis for childless singles is based on the elasticities suggested by Chetty, Guren, Manoli and Weber (2013): an intensive-margin elasticity of 0.33, and a participation elasticity that equals 0.25 on average. For single parents, we also use an intensive-margin elasticity of 0.33, and an average participation elasticity of 0.58, as estimated by Bastian (2020) based on the 1975 EITC reform.[39] For both groups, we assume, in line with the empirical evidence, that participation elasticities decline with income (see part C of the Supplementary Material for details).

[36] For the benchmark analysis reported below, we estimated the income distribution based on the set of single parents with any number of children ($N = 1,494$). As a robustness check, we estimate this distribution for the smaller sample of single parents with exactly two children ($N = 453$). In part C of the Supplementary Material, we show that this is inconsequential for the results of our analysis.

[37] Eligibility for EITC, AFDC and SNAP also involved assets and capital income tests. According to CPS data, more than 90% of single parents satisfied these tests. In our main analysis, we therefore ignore them (i.e., we assume that all single parents satisfy the tests). Part C of the Supplementary Material provides further details and shows that our qualitative results remain unchanged if we explicitly account for the assets and capital income tests.

[38] Our results are not affected if we instead use parametric estimates of the income distribution, see part C of the Supplementary Material for further details.

[39] Kleven (2019) finds an extensive margin elasticity close to zero for single parents based on later EITC reforms. We consider this case in our sensitivity analysis below.

5.3 Empirical results

In the following, we present our benchmark calibrations of the revenue functions $\mathcal{R}_{sp}(y)$ and $\mathcal{R}_{cs}(y)$ for *s*ingle *p*arents and *c*hildless *s*ingles, respectively. We then use these functions to investigate, first, whether the US tax-and-transfer system was Pareto-efficient prior to the EITC introduction and, second, whether the 1974 EITC reform for single parents was a reform in a Pareto-improving direction. Third, we characterize the best available reform direction at the time, i.e., the revenue-maximizing Pareto improvement.

Was the 1974 US tax-and-transfer system Pareto-efficient? Figure 5 plots the revenue functions $y \mapsto \mathcal{R}_{sp}(y)$ and $y \mapsto \mathcal{R}_{cs}(y)$ for our benchmark calibration of the 1974 US tax system. Specifically, the solid blue line depicts the revenue function $\mathcal{R}_{sp}(y)$ for single parents, while the teal line depicts the revenue function $\mathcal{R}_{cs}(y)$ for childless singles.

For single parents, function $\mathcal{R}_{sp}(y)$ does not satisfy the necessary conditions for Pareto efficiency in Theorem 1. First, it attains negative values for incomes between approximately $1,500$ and $5,400$ USD. This implies that one-bracket tax cuts in this income range would have been self-financing and Pareto-improving. Second, $y \mapsto \mathcal{R}_{sp}(y)$ is increasing in the income range between $5,000$ and $6,000$ USD, thereby violating the monotonicity condition. Hence, there was room for Pareto-improving two-bracket tax cuts, resembling the EITC.

For childless singles, the revenue response function $y \mapsto \mathcal{R}_{cs}(y)$ is throughout between 0 and 1, so that there was no scope for a Pareto-improving reform involving only a single bracket. By contrast, the monotonicity condition on $y \mapsto \mathcal{R}_{cs}(y)$ is violated as the teal line is slightly increasing in the range between $2,000$ and $3,000$ USD. Again, this indicates the possibility of Pareto-improving two-bracket reforms. That said, the visual impression is that the scope for such a Pareto improvement was more limited for childless singles than for single parents. Below, we confirm this conjecture using the inefficiency measure introduced in Section 4 above.

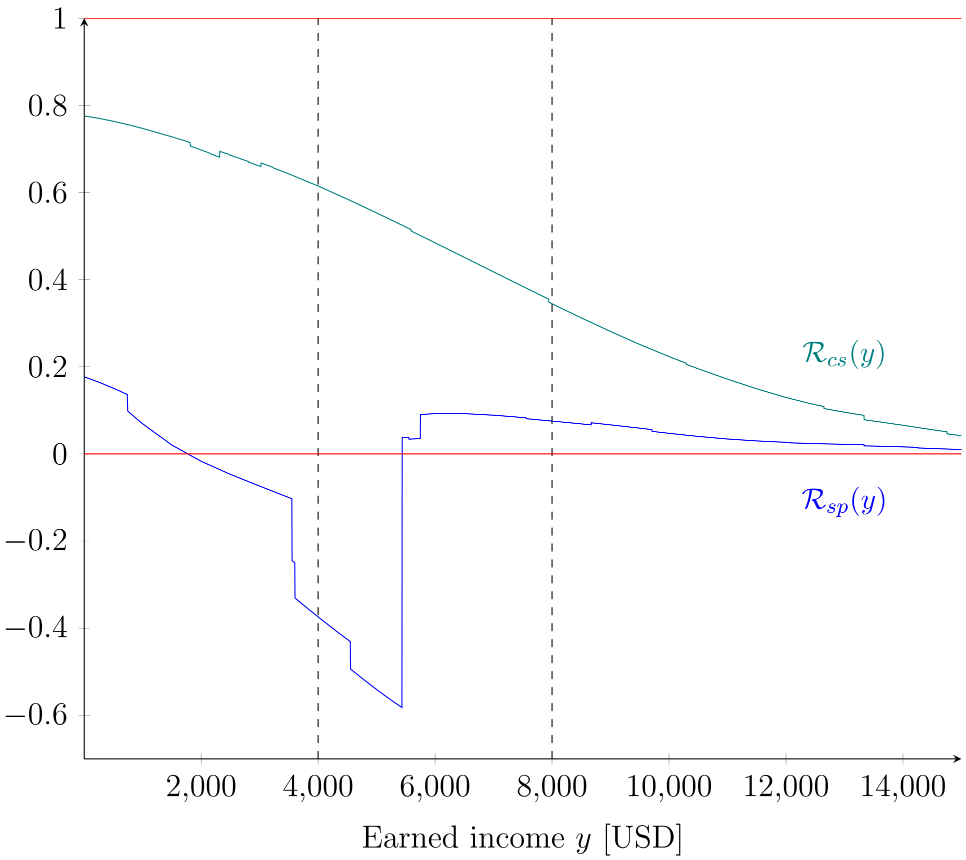


Figure 5: Pareto test of the 1974 US tax-transfer system.

Notes: Figure 5 shows the revenue functions for single parents $y \mapsto \mathcal{R}_{sp}(y)$ (blue) and childless singles $y \mapsto \mathcal{R}_{cs}(y)$ (teal) in 1974 as functions of earned income for our benchmark calibration: intensive-margin elasticities of 0.33 for both groups, average participation elasticities of 0.58 for single parents and 0.25 for childless singles. The vertical dashed lines mark the endpoints of the phase-in range at $4,000$ USD and the phase-out range at $8,000$ USD of the 1974 EITC.

Source: Authors' calculations (see part C of the Supplementary Material for details).

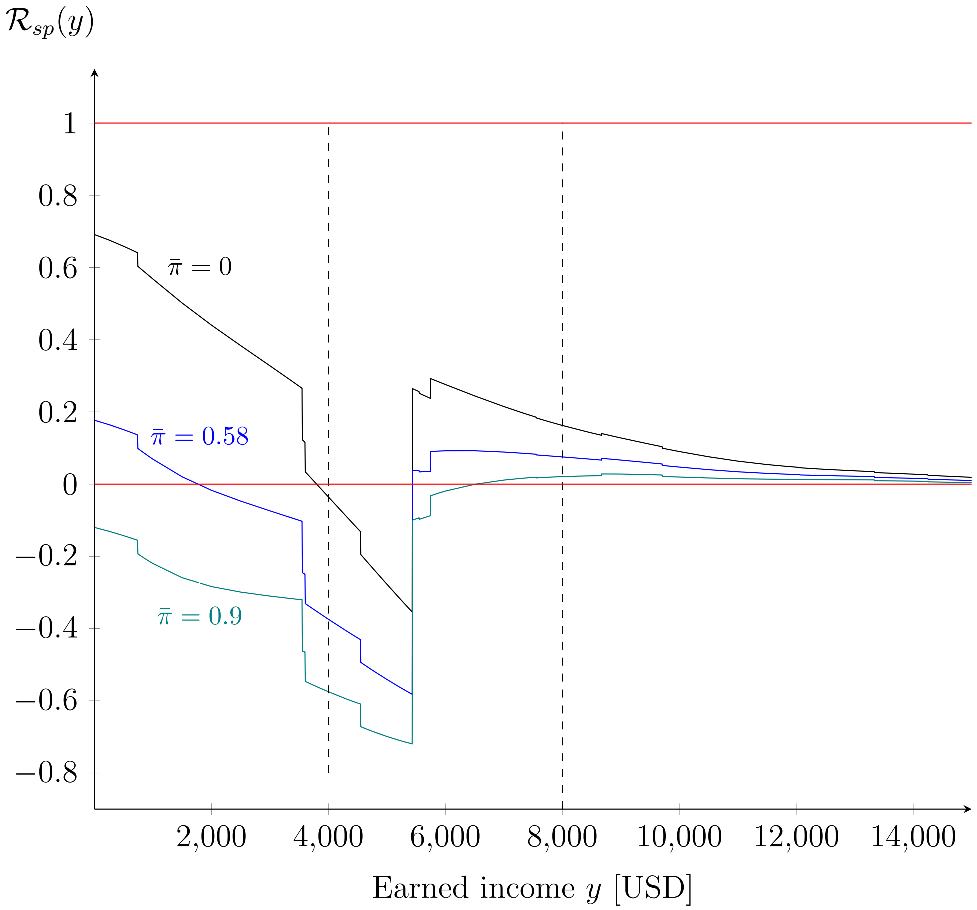


Figure 6: Pareto test of 1974 US tax system, different participation elasticities.

Notes: Figure 6 shows the revenue function $\mathcal{R}_{sp}(y)$ for single parents in 1974, assuming an average participation elasticity of 0.58 (blue line, benchmark), a higher participation elasticity of 0.9 (teal line) and a case without extensive-margin responses (black line). The intensive-margin elasticity is held at the benchmark level of 0.33. The dashed vertical lines mark the endpoints of the phase-in range at $4,000$ USD and the phase-out range at $8,000$ USD of the 1975 EITC.

Source: Authors' calculations (see part C of the Supplementary Material for details).

Sensitivity analysis. The finding that it was possible to realize a Pareto improvement by means of a two-bracket tax cut for single parents is robust in various dimensions. For instance, Figure 6 explores alternative assumptions about behavioral responses at the extensive margin: It shows the revenue function $y \mapsto \mathcal{R}_{sp}(y)$ for our benchmark calibration with an average participation elasticity of 0.58 (blue line), for a case with a larger participation elasticity of 0.9 (teal line), and for the limit case with a participation elasticity of zero (black line). The figure suggests that the scope for Pareto-improving reforms is larger, the more strongly labor supply responds at the extensive margin. But this scope does not vanish if the participation elasticity is zero. This observation is interesting in the light of the discussion about the EITC from an optimal tax perspective, where positive extensive-margin elasticities are frequently considered necessary for the desirability of an EITC, see, e.g., Saez (2002) or Hansen (2021). As we show here, with a tax reform perspective applied to the tax and transfer system as of 1974, the introduction of the EITC can be rationalized even when there are no behavioral responses at the extensive margin.[40]

Was the 1975 EITC reform Pareto-improving? Figure 5 shows that it was possible to Pareto-improve the US tax-transfer system by a two-bracket tax cut, i.e., by the introduction of *some EITC*. A separate question is whether the EITC reform that *actually took place* went into a Pareto-improving direction. To answer this question, we make use of the conditions in Proposition 2 which combine an estimate of the revenue effects with information on the reform direction, referred to as $\tilde{h}_{75}$ below. Specifically, the 1975 EITC reform reduced marginal taxes at all incomes below $4,000$ USD by 10 percentage points, and increased marginal taxes at all incomes between $4,000$ and $8,000$ USD by the same amount. It did not increase tax liabilities at any income level. This reform had a Pareto-improving

[40]Part C.2 of the Supplementary Material contains further robustness checks that consider, e.g., other assumptions about intensive-margin elasticities, income effects, single parents with one or three children, alternative estimates of the income distribution, or other representations of the tax-transfer system.

direction if the condition

$$R_\tau(0, \tilde{h}_{75}) = -\int_0^{4,000} \mathcal{R}_{sp}(y)dy + \int_{4,000}^{8,000} \mathcal{R}_{sp}(y)dy \quad \geq \quad 0 \tag{12}$$

is satisfied. Whether this inequality holds depends on the details of the calibration. For the benchmark case with an intensive-margin elasticity of 0.33 and an average participation elasticity of 0.58, the reform was not Pareto-improving. For a participation elasticity above 0.84, by contrast, it was Pareto-improving.[41]

The revenue-maximizing tax cut. What would have been the revenue-maximizing tax cut in 1975 and how does it relate to the reform that actually was taken? Under our benchmark calibration, the former is a two-bracket tax cut that reduces marginal tax rates between $1,248$ and $5,748$ USD, and increases marginal taxes between $5,748$ and $10,248$ USD. Thus, the optimal reform would have been a version of the EITC that involved a wider range of incomes, and also higher incomes than the actual 1975 EITC. With the revenue gain from this reform, it would have been possible to pay an additional lump-sum transfer of 12.6 USD per percentage-point change in marginal taxes to each single parent (corresponding to 71 USD in 2021). Figure 7 illustrates this reform and its revenue implications graphically. For childless singles, the corresponding number is much smaller, namely 1 cent per percentage point change in marginal taxes. This confirms the conjecture above that the inefficiency in the tax and transfer system for childless singles was orders of magnitude smaller than the one for single parents.

Did subsequent reforms improve the EITC? Since its introduction, the EITC was repeatedly reformed and expanded in two major ways. First, the range of eligible incomes was enlarged in several steps, starting with the 1979 reform. Second, benefits were made dependent on family size, with larger benefits for larger families; e.g., in 1991 and 2009. Did these reforms Pareto-improve the EITC or, put differently, was there progress in US tax policy for people with low

[41] Some studies find participation elasticities in this range, see, e.g., Meyer and Rosenbaum (2001).

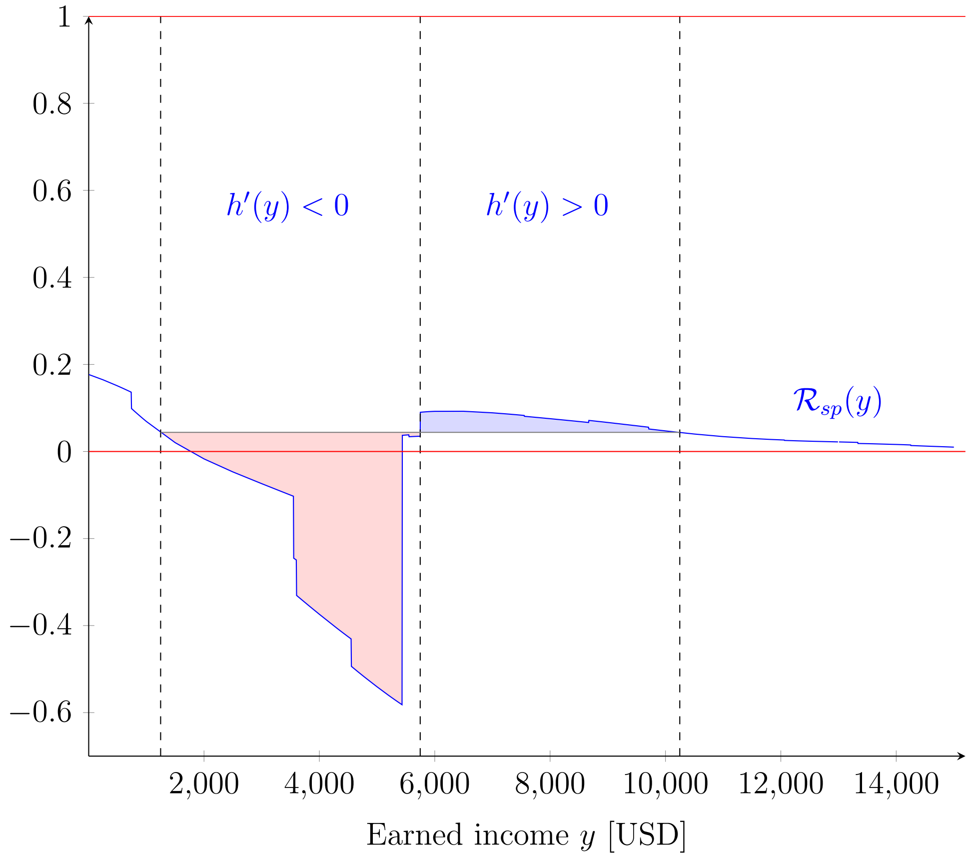


Figure 7: The revenue-maximizing tax cut for single parents.

Notes: Figure 7 shows the revenue function $y \mapsto \mathcal{R}_{sp}(y)$ for single parents given our benchmark calibration. The revenue-maximizing tax cut would have reduced marginal taxes between $1,248$ (first dashed line) and $5,748$ USD (second dashed line) and increased them between $5,748$ and $10,248$ USD (third dashed line). The sum of both shaded areas represents the available revenue gain per single-parent family.

Source: Authors' calculations (see part C of the Supplementary Material for details).

incomes? Our approach to study Pareto-improving tax reforms can also be used to answer these questions. The following paragraphs summarize our findings; a more detailed analysis can be found in part C of the Supplementary Material.

First, we study the efficiency of the US tax-transfer system between the EITC introduction in 1975 and the first EITC expansion in 1979. As of 1974, the revenue-maximizing tax cut would have reduced marginal taxes between $1,248$ and $5,748$ USD. The 1975 EITC only reduced marginal taxes below $4,000$ USD, however, and increased them between $4,000$ and $8,000$ USD. Thus, the actual reform aggravated inefficiencies at the bottom of the phase-out range. Correspondingly, the 1975 US tax-transfer system remained Pareto-inefficient, and there was further scope for Pareto-improving reforms at higher incomes. In 1979, the US government indeed extended the EITC to higher income levels – by means of a two-bracket tax cut affecting incomes between $4,000$ and $10,000$ USD. We demonstrate that, based on our benchmark calibration, the 1979 EITC reform actually had a Pareto-improving direction. Thus, while the initial version of the EITC was suboptimal, its design was improved subsequently.

Second, we look into the desirability of making the EITC provisions dependent on the number of children. We find that, for each group, the introduction of an EITC schedule in 1974 was Pareto-improving. The scope for Pareto improvements was even larger with differentiated schedules that were more generous for larger families. The 1974 EITC did not condition on the number of children. More than a decade later, the US government improved the design of the EITC in this dimension, introducing more generous tax credits for families with two and more children in 1991, and for families with three and more children in 2009.

6 Concluding Remarks

A key lesson from this paper is that tax reforms with two brackets – one in which tax rates are lowered, and one in which tax rates are increased – deserve particular attention. Our theoretical results show that such reforms can make every one better off, even if no simple one-bracket tax reform can. Moreover, a tax system is Pareto-efficient if there is no Pareto improvement in the class of tax reforms that affect at most two brackets. Our study of the EITC shows that such reforms have also been successfully used in practice.[42] Finally, we show that sufficient-statistics formulas can be used to identify Pareto-improving two-bracket reforms in practice.

References

Akerlof, G.A., "The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning," *American Economic Review*, 1978, *68* (1), 8–19.

Badel, A. and M. Huggett, "The sufficient statistic approach: Predicting the top of the Laffer curve," *Journal of Monetary Economics*, 2017, *87*, 1–12.

Bargain, O., M. Dolls, D. Neumann, A. Peichl, and S. Siegloch, "Tax-Benefit Systems in Europe and the US: Between Equity and Efficiency," *IZA Discussion Papers 5440*, 2011.

___ , ___ , ___ , ___ , and ___ , "Comparing inequality aversion across countries when labor supply responses differ," *International Tax and Public Finance*, Oct 2014, *21* (5), 845–873.

Bastani, S., S. Blomquist, and L. Micheletto, "Pareto efficient income taxation without single-crossing," *Social Choice and Welfare*, 2020.

Bastian, J.E., "The Rise of Working Mothers and the 1975 Earned Income Tax Credit," *American Economic Journal: Economic Policy*, 2020, *12* (3), 44–75.

___ and M.R. Jones, "Do EITC Expansions Pay for Themselves? Effects on Tax Revenue and Public Assistance Spending," *Journal of Public Economics*, 2021, *196*, 104355.

[42] Interestingly, the US tax and transfer system of 2018 has a pattern of inefficiencies that is qualitatively similar to those in the mid 70s, but quantitatively less significant, see part G of the Supplementary Material.

Bierbrauer, F.J. and P.C. Boyer, "The Pareto-frontier in a simple Mirrleesian model of income taxation," *Annals of Economics and Statistics*, 2014, (113/114), 185–206.

___ **and** ___ , "Politically feasible reforms of non-linear tax systems," *CEPR Discussion Paper 13059*, 2018.

___ **,** ___ **, and A. Peichl**, "Politically feasible reforms of nonlinear tax systems," *American Economic Review*, 2021, *111* (1), 153–91.

Blundell, R., "Earned income tax credit policies: Impact and optimality: The Adam Smith Lecture, 2005," *Labour Economics*, 2006, *13* (4), 423–443.

___ **and A. Shephard**, "Employment, hours of work and the optimal taxation of low-income families," *Review of Economic Studies*, 2012, *79* (2), 481–510.

___ **, M. Brewer, P. Haan, and A. Shephard**, "Optimal income taxation of lone mothers: an empirical comparison of the UK and Germany," *Economic Journal*, 2009, *119*, F101–F121.

Boadway, R. and M. Keen, "Redistribution," in A.B. Atkinson and F. Bourguignon, eds., *Handbook of Income Distribution*, Vol. 1, Elsevier, 2000, pp. 677–789.

Bourguignon, F. and A. Spadaro, "Tax-Benefit Revealed Social Preferences," *Journal of Economic Inequality*, 2012, *10* (1), 75–108.

Brett, C. and J.A. Weymark, "Voting over selfishly optimal nonlinear income tax schedules," *Games and Economic Behavior*, 2017, *101*, 172–188.

Brito, D.L., J.H. Hamilton, S.M Slutsky, and J.E. Stiglitz, "Pareto Efficient Tax Structures," *Oxford Economic Papers*, 1990, *42* (1), 61–77.

Cesarini, D., E. Lindqvist, M.J. Notowidigdo, and R. Östling, "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries," *American Economic Review*, 2017, *107* (12), 3917–46.

Chetty, R., "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 2009, *1* (1), 451–488.

——, **A. Guren, D. Manoli, and A. Weber**, "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," *NBER Macroeconomics Annual*, 2013, *27* (1), 1–56.

Choné, P. and G. Laroque, "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 2010, *100* (5), 2532–2547.

—— **and** ——, "Optimal taxation in the extensive model," *Journal of Economic Theory*, 2011, *146* (2), 425–453.

Christiansen, V. and E. Jansen, "Implicit Social Preferences in the Norwegian System of Indirect Taxation," *Journal of Public Economics*, 1978, *10* (2), 217–245.

Diamond, P.A., "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 1998, *88*, 83–95.

Eissa, N. and H.W. Hoynes, "Taxes and the labor market participation of married couples: the Earned Income Tax Credit," *Journal of Public Economics*, 2004, *88*, 1931–1958.

—— **and J.B. Liebman**, "Labor Supply Response to the Earned Income Tax Credit," *Quarterly Journal of Economics*, 1996, *111* (2), 605–637.

Feldstein, M., "On the theory of tax reform," *Journal of Public Economics*, 1976, *6* (1-2), 77–104.

Giannarelli, L., *An Analyst's Guide to TRIM2: The Transfer Income Model, Version 2*, Washington: The Urban Institute Press, 1992.

Golosov, M., A. Shourideh, M. Troshkin, and A. Tsyvinski, "Optimal pension systems with simple instruments," *American Economic Review: Papers and Proceedings*, 2013, *103*, 502–507.

——, **A. Tsyvinski, and N. Werquin**, "A Variational Approach to the Analysis of Tax Systems," *NBER Working Paper 20780*, 2014.

Guesnerie, R., *A Contribution to the Pure Theory of Taxation*, Cambridge University Press, 1995.

Hansen, E., "Optimal Income Taxation with Labor Supply Responses at Two Margins: When Is An Earned Income Tax Credit Optimal?," *Journal of Public Economics*, 2021, *195*, 104365.

Hendren, N., "Measuring Economic Efficiency Using Inverse Optimum Weights," *Journal of Public Economics*, 2020, *187*, 104198.

Holtz-Eakin, D., D. Joulfaian, and H.S. Rosen, "The Carnegie Conjecture: Some Empirical Evidence," *Quarterly Journal of Economics*, 1993, *108* (2), 413–435.

Hotz, V.J. and J.K. Scholz, "The Earned Income Tax Credit," in R.A. Moffitt, ed., *Means-Tested Transfer Programs in the United States*, University of Chicago Press, 2003, pp. 141–198.

Hoynes, H., "The Earned Income Tax Credit," *The Annals of the American Academy of Political and Social Science*, 2019, *686* (1), 180–203.

Imbens, G.W., D.B. Rubin, and B.I. Sacerdote, "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, 2001, *91* (4), 778–794.

Jacobs, B., E. Jongen, and F. Zoutman, "Revealed social preferences of Dutch political parties," *Journal of Public Economics*, 2017, *156*, 81–100.

Jacquet, L. and E. Lehmann, "Optimal Income Taxation with Composition Effects," *Journal of the European Economic Association*, 2021, *19* (2), 1299–1341.

——, ——, **and B. Van der Linden**, "Optimal redistributive taxation with both extensive and intensive responses," *Journal of Economic Theory*, 2013, *148* (5), 1770–1805.

Juhn, C., K.M. Murphy, and R.H. Topel, "Why Has the Natural Rate of Unemployment Increased over Time?," *Brookings Papers on Economic Activity*, 1991, (2), 75–142.

——, ——, **and** —— , "Current Unemployment, Historically Contemplated," *Brookings Papers on Economic Activity*, 2002, (1), 79–136.

Kleven, H., "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 2014, *28* (4), 77–98.

—— , "The EITC and the Extensive Margin: A Reappraisal," NBER Working Paper No. 26405, 2019.

—— , "Sufficient Statistics Revisited," *Annual Review of Economics*, 2021, *13*, 515–538.

Koehne, S. and D. Sachs, "Pareto-efficient Tax Deductions," CESifo Working Paper No. 7744, 2019.

Lockwood, B.B. and M. Weinzierl, "Positive and normative judgments implicit in U.S. tax policy, and the costs of unequal growth and recessions," *Journal of Monetary Economics*, 2016, *14-119* (77), 30–47.

Lorenz, N. and D. Sachs, "Identifying Laffer Bounds: A Sufficient-Statistics Approach with an Application to Germany," *Scandinavian Journal of Economics*, 2016, *118* (4), 646–665.

Meghir, C. and D. Phillips, "Labour Supply and Taxes," in J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba, eds., *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press, 2010, pp. 202–274.

Meyer, B.D. and D.T. Rosenbaum, "Welfare, the Earned Income Tax Credit and the Labor Supply of Single Mothers," *Quarterly Journal of Economics*, 10 2001, *116* (3), 1063–1114.

Michau, J.-B., "Optimal redistribution: A life-cycle perspective," *Journal of Public Economics*, 2014, *111*, 1–16.

Milgrom, P. and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 2002, *70* (2), 583–601.

Mirrlees, J., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38*, 175–208.

Moffitt, R.A., "The Negative Income Tax and the Evolution of U.S. Welfare Policy," *Journal of Economic Perspectives*, 2003, *17* (3), 119–140.

Nichols, A. and J. Rothstein, "The Earned Income Tax Credit (EITC)," *NBER Working Paper Series*, 2015.

Piketty, T., "La redistribution fiscale face au chômage," *Revue française d'économie*, 1997, *12*, 157–201.

—— **and E. Saez**, "Optimal Labor Income Taxation," in A.J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, eds., *Handbook of Public Economics*, Vol. 5, Elsevier, 2013, pp. 391–474.

Saez, E., "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68*, 205–229.

——, "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 2002, *117* (3), 1039–1073.

——— **and S. Stantcheva**, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 2016, *106* (1), 24–45.

——— **, J. Slemrod, and S.H. Giertz**, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 2012, *50* (1), 3–50.

Scheuer, F., "Entrepreneurial taxation with endogenous entry," *American Economic Journal: Economic Policy*, 2014, *6* (2), 126–63.

Shourideh, A. and M. Troshkin, "Incentives and Efficiency of Pension Systems," *Working Paper*, 2017.

Stantcheva, S., "Optimal income taxation with adverse selection in the labour market," *Review of Economic Studies*, 2014, *81* (3), 1296–1329.

Stiglitz, J., "Self-Selection and Pareto-Efficient Taxation," *Journal of Public Economics*, 1982, *17*, 213–240.

———, "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," in A.J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, Vol. 2, Elsevier, 1987, pp. 991–1042.

Ventry, D.J., "The Collision of Tax and Welfare Politics: The Political History of the Earned Income Tax Credit, 1969–99," *National Tax Journal*, 2000, *53* (4), 983–1026.

Werning, I., "Pareto Efficient Income Taxation," Working paper 2007.

Weymark, J.A., "Undominated directions of tax reform," *Journal of Public Economics*, 1981, *16* (3), 343–369.

Zorn, M.A., "Characterization of analytic functions in Banach spaces," *Annals of Mathematics*, 1945, *46* (4), 585–593.

Appendix

A Proofs

A.1 Proof of Theorem 1 and Proposition 1.

To prove Theorem 1, we proceed in two steps. We first show that the non-existence of a Pareto-improving one-bracket reform implies that $\mathcal{R}(y) \in [0,1]$ for all y. We then show that the non-existence of a Pareto-improving two-bracket reform implies that $y \mapsto \mathcal{R}(y)$ is non-increasing.

Reforms with one bracket. Adapting inequality (5) to the case of a one-bracket reform we find that a small reform with $\tau > 0$ that increases marginal tax rates from the status quo is Pareto-improving if, for some $\ell > 0$,

$$R^s_\tau(0, \ell, \hat{y}) - \ell > 0 \;, \tag{13}$$

i.e., if marginal revenue gains are so large that even those agents are made better off whose tax bill increases by the maximal amount of $\max_y \; h^s(y) = \ell$. For a one-bracket reform with $\tau < 0$, we have to compare the derivative $R^s_\tau(0, \ell, \hat{y})$ with $\min_y \; h^s(y) = 0$, see Footnote 19. Consequently, a small one-bracket reform that reduces marginal tax rates is Pareto-improving if

$$R^s_\tau(0, \ell, \hat{y}) < 0 \;, \tag{14}$$

so that a tax cut leads to larger tax revenues, a logic familiar from analyses of the Laffer curve.

Below we exploit the fact that a one-bracket reform on a bracket of length zero does not affect tax revenue, i.e., $R^s_\tau(0, 0, \hat{y}) = 0$ for any $\hat{y}$. To see this, recall that the new tax schedule satisfies $T_1(y) = T_0(y)$ for any $y \leq \hat{y}$, and $T_1(y) = T_0(y) + \tau \, \ell$ for any $y \geq \hat{y} + \ell$. For a one-bracket reform on a bracket of length $\ell = 0$, the new tax schedule is thus identical to the status quo tax schedule, $T_1(y) = T_0(y)$ for all y, independent of the step size τ. Hence, a variation in τ neither affects the budget set C_1 nor individual behavior y^*, nor tax revenue $R^s(\tau, 0, \hat{y})$.

Lemma 1

(i) If $\mathcal{R}(\hat{y}) - 1 > 0$ for some $\hat{y} \in \mathcal{Y}$, there exists a Pareto-improving one-bracket reform with $\tau > 0$ and $\ell > 0$.

(ii) If $\mathcal{R}(\hat{y}) < 0$ for some $\hat{y} \in \mathcal{Y}$, there exists a Pareto-improving one-bracket reform with $\tau < 0$ and $\ell > 0$.

(iii) If there is no Pareto-improving reform, then $\mathcal{R}(\hat{y}) \in [0, 1]$ for all $\hat{y}$.

Proof of Lemma 1. As explained above, we have $R^s_\tau(0, 0, \hat{y}) = 0$. If $\mathcal{R}(\hat{y}) = R^s_{\tau\ell}(0, 0, \hat{y}) > 1$, this implies that $R^s_\tau(0, \ell, \hat{y}) - \ell$ turns positive if, starting from $\ell = 0$, the length of the interval is slightly increased. This proves (i). Analogously, if $\mathcal{R}(\hat{y}) = R^s_{\tau\ell}(0, 0, \hat{y}) < 0$, this implies that $R^s_\tau(0, \ell, \hat{y})$ turns negative if, starting from $\ell = 0$, the length of the interval is slightly increased. This proves (ii). Thus, necessary conditions for the non-existence of a Pareto-improving one-bracket reform are $\mathcal{R}(\hat{y}) \leq 1$ and $\mathcal{R}(\hat{y}) \geq 0$. This proves (iii).

Reforms with two brackets. Lemma 1 above gives necessary conditions for the existence of Pareto-improving reforms with a single bracket. The following Lemma gives the analogue for the case of two-bracket reforms. In particular, we show that, if $y \mapsto \mathcal{R}(y)$ is increasing, the combination of two reforms – each of which would not be Pareto-improving on a stand alone basis – yields a Pareto improvement. For this purpose, we denote by $R(\tau, h_2)$ the change in tax revenue due to a joint reform with two brackets, where $h_2 = h^s_1 + h^s_2$ is composed of two single bracket reforms.

Proof of Proposition 1. Fix two income levels y_1 and y_2 such that $y_2 > y_1$ and $\mathcal{R}(y_2) > \mathcal{R}(y_1)$. We now construct a Pareto-improving two-bracket reform with the parameters $\{(y_1, \tau_1, \ell_1, y_2, \tau_2, \ell_2, \tau, \ell)\}$. In particular, let $\tau_1 < 0$, $\tau_2 > 0$, and $\tau_1\, \ell_1 + \tau_2\, \ell_2 = 0 > \tau_1\, \ell_1$. This implies that $\max_y h(y) = 0$. By the linearity of the

Gateaux differential,[43] we moreover find that

$$
\begin{aligned}
R_\tau(0, h_2) &= \tau_1 R^s_\tau(0, \ell\,\ell_1, y_1) + \tau_2 R^s_\tau(0, \ell\,\ell_2, y_2)\ , \text{ and} \\
R_{\tau\ell}(0, h_2) &= \tau_1 \ell_1 R^s_{\tau\ell}(0, 0, y_1) + \tau_2 \ell_2 R^s_{\tau\ell}(0, 0, y_2) \\
&= \tau_2 \ell_2 \left[\mathcal{R}(y_2) - \mathcal{R}(y_1)\right] > 0\ .
\end{aligned}
$$

Hence, there exists $\hat{\ell} > 0$ such that $R_\tau(0, h_2) - \max_y h(y) > 0$ for all $\ell \in (0, \hat{\ell})$. Finally, by equation (4), this implies that for a reform as constructed above with $\ell \in (0, \hat{\ell})$, a small increase in τ is Pareto-improving.

Suppose that $\mathcal{R}(y_1)$ and $\mathcal{R}(y_2)$ are between 0 and 1. Then, there is no Pareto-improving one-bracket reform for incomes close to y_1 or close to y_2. If $\mathcal{R}(y_1) < \mathcal{R}(y_2)$, however, there is still scope for a Pareto improvement that involves two brackets.

A.2 Proof of Theorem 2

A reform with an arbitrary number m of brackets can be characterized as a collection

$$
\{(y_k, \tau\ \tau_k, \ell\ \ell_k)\}_{k=1}^m
$$

of one-bracket reforms, where the marginal tax in the kth bracket is changed by $\tau\ \tau_k$ and length of the kth bracket is given by $\ell\ \ell_k$. As before, the parameters (τ, ℓ) determine the size of the reform and the overall revenue is denoted by $R(\tau, h_m)$. The following lemma states sufficient conditions for the existence of a Pareto-improving reform with m brackets.

[43] Gateaux differentials are not linear in general. To clarify the conditions under which they are, for $0 < \bar{y}, a < \infty$, let $\tau \in [-a, a]$ and $h \in \mathcal{H} := (\mathcal{C}[0, \bar{y}], \|\cdot\|_\infty)$, where $\|\cdot\|_\infty$ denotes the sup norm. We define the operator

$$
\mathcal{R}:\ \ \mathcal{H} \to \mathcal{K}:\ \ h \mapsto \mathcal{R}(\tau, h),
$$

where $\mathcal{K} := (\mathcal{C}_b([-a, a]), \|\cdot\|_\infty)$ and $\mathcal{C}_b([-a, a])$ denotes the set of bounded continuous real functions defined on $[-a, a]$. Note that $\mathcal{H}$ and $\mathcal{K}$ are Banach spaces. In this setting, the Gateaux differential of $R_\tau(\tau, \cdot)$ is linear (Zorn (1945); Theorem 2.3).

Lemma 2 *Consider a collection* $\{(y_k, \tau\ \tau_k, \ell\ \ell_k)\}_{k=1}^m$ *of simple reforms. Let* $\tau_0\ l_0 = 0$. *There is a reform* (τ, h_m) *with* $\tau > 0$ *and* $\ell > 0$ *that is Pareto-improving if*

$$\sum_{k=1}^{m} \tau_k\ \ell_k\ \mathcal{R}(y_k) - \max_{j \in \{0,1,\ldots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k \quad > \quad 0\ .$$

Proof The linearity of the Gateaux differential implies that

$$\begin{aligned} R_\tau(0, h_m) &= \sum_{k=1}^{m} \tau_k R_\tau^s(0, \ell\ \ell_k, y_k)\ , \text{ and} \\ R_{\tau\ell}(0, h_m) &= \sum_{k=1}^{m} \tau_k\ \ell_k\ \mathcal{R}(y_k)\ . \end{aligned}$$

Moreover,

$$\max_y\ h_m(y) \;=\; \ell \max_{j \in \{0,1,\ldots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k\ .$$

As shown above, $R_\tau(0, h_m)$ equals zero for a reform with $\ell = 0$ such that all brackets have length zero. Hence, if the condition in the lemma is satisfied, there exists a $\hat{\ell} > 0$ such that $R_\tau(0, h_m) - \max_y\ h_m(y) > 0$ for all $\ell \in (0, \hat{\ell})$. By equation (4), this implies that for such an m-bracket reform (τ, h_m) with $\ell \in (0, \hat{\ell})$, a small increase in the step size τ is Pareto-improving.

Lemma 2 states sufficient conditions for the existence of Pareto-improving reforms. If we limit attention to small reforms these conditions are also necessary, i.e., if they do not hold there is no small reform that is Pareto-improving. The following lemma shows that, if the conditions in Theorem 2 hold, the condition in Lemma 2 is violated for any collection of m single bracket reforms. Consequently, there is no small Pareto-improving m-bracket reform.

Lemma 3 *Suppose that the function* $y \mapsto \mathcal{R}(y)$ *is bounded from below by* 0, *bounded from above by 1 and non-increasing. Let* $\tau_0\ l_0 = 0$. *Then,*

$$\sum_{k=1}^{m} \tau_k\ \ell_k\ \mathcal{R}(y_k) - \max_{j \in \{0,1,\ldots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k \quad \leq \quad 0 \tag{15}$$

for any collection $\{(y_k, \tau_j\ \tau_k, \ell_j\ \ell_k)\}_{k=1}^m$, *and for any* $m \geq 1$.

Proof of Lemma 3. Let j^* be a bracket in which the function h_m achieves a maximum, $j^* := \text{argmax}_j \quad \sum_{k=0}^{j} \tau_k \, \ell_k$. Note that this implies that $\sum_{k=z}^{j^*} \tau_k \, \ell_k \geq 0$ for any $z \in \{0, \ldots, j^*\}$ and $\sum_{k=j^*+1}^{z} \tau_k \, \ell_k \leq 0$ for any $z \in \{j^*+1, \ldots, m\}$; otherwise j^* would not be a maximizer.

Step 1. We verify the following claim: Suppose that $j^* > 0$ and that

$$\sum_{k=z}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k) \quad \leq \quad \mathcal{R}(y_z) \sum_{k=z}^{j^*} \tau_k \, \ell_k \tag{16}$$

holds for some $z \in \{1, \ldots, j^*\}$. Then, if $z > 1$, we also have

$$\begin{aligned}
\sum_{k=z-1}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k) \quad &= \quad \tau_{z-1} \, l_{z-1} \, \mathcal{R}(y_{z-1}) + \sum_{k=z}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k) \\
&\leq \quad \tau_{z-1} \, l_{z-1} \, \mathcal{R}(y_{z-1}) + \underbrace{\mathcal{R}(y_z)}_{\leq \mathcal{R}(y_{z-1})} \underbrace{\sum_{k=z}^{j^*} \tau_k \, \ell_k}_{\geq 0} \\
&\leq \quad \mathcal{R}(y_{z-1}) \sum_{k=z-1}^{j^*} \tau_k \, \ell_k \, .
\end{aligned}$$

Condition (16) is obviously satisfied for $z = j^*$. Hence, a repeated application of the preceding argument yields

$$\sum_{k=1}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k) \quad \leq \quad \underbrace{\mathcal{R}(y_1)}_{\in[0,1]} \underbrace{\sum_{k=1}^{j^*} \tau_k \, \ell_k}_{\geq 0} \quad \leq \quad \sum_{k=1}^{j^*} \tau_k \, \ell_k \quad = \quad \sum_{k=0}^{j^*} \tau_k \, \ell_k \, . \tag{17}$$

Step 2. An analogous argument implies that

$$\sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, \mathcal{R}(y_k) \quad \leq \quad \mathcal{R}(y_m) \underbrace{\sum_{k=j^*+1}^{m} \tau_k \, \ell_k}_{\leq 0} \quad \leq \quad 0 \, . \tag{18}$$

Step 3. Together (17) and (18) imply that, if $j^* \in \{1, \ldots, m-1\}$,

$$R_{\tau\ell}(0, h_m) \quad = \quad \sum_{k=1}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k) + \sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, \mathcal{R}(y_k) \quad \leq \quad \sum_{k=0}^{j^*} \tau_k \, \ell_k \, , \tag{19}$$

which proves (15). Note that the cases $j^* = 0$ and $j^* = m$ are also covered. With $j^* = 0$, $\sum_{k=1}^{j^*} \tau_k \, \ell_k \, \mathcal{R}(y_k)$ does not enter the chain of inequalities and (18) directly implies (15). With $j^* = m$, $\sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, \mathcal{R}(y_k)$ does not enter and (17) directly implies (15).

A.3 Proof of Corollary 1 and Proposition 2

Take any continuous reform direction h on $[0, \bar{y}]$. We approximate h with a piecewise linear reform direction h_m that involves m one-bracket reforms $(\tau_k, h_k^s)_{k=1}^m$, so that $h_m(y) = \sum_{k=1}^m \tau_k \; h_k^s(y)$. Throughout, we let $\ell = 1$ and divide the domain $[0, \bar{y}]$ into m brackets of equal length $\ell_k = \frac{1}{m}\bar{y}$, starting at incomes $y_1 = 0$, $y_2 = \frac{1}{m}\bar{y}$, …, and $y_m = \frac{m-1}{m}\bar{y}$. Thus, we have m adjacent brackets – a special case of our general formalism, which also allows for gaps between the brackets where marginal tax rates change. For any k, we then let

$$\tau_k = \frac{h(y_{k+1}) - h(y_k)}{\ell_k} \; , \quad \text{where we set} \quad y_{m+1} = \bar{y}.$$

This yields an approximation of h by a piecewise linear function. The construction is illustrated in Figure 8. By choosing m sufficiently large, the piecewise linear function h_m approximates h in the sense that, for any $\varepsilon > 0$, there exists $\hat{m}(\varepsilon)$ so that for any $m > \hat{m}(\varepsilon)$,

$$\sup_{y \in \mathcal{Y}} \; | \; h(y) - h_m(y) \; | \quad < \quad \varepsilon \; .$$

For later reference, we note that this implies in particular that, for any y^* that maximizes $h(y)$ over $\mathcal{Y}$, we have

$$h(y^*) - h_m(y^*) < \varepsilon \; . \tag{20}$$

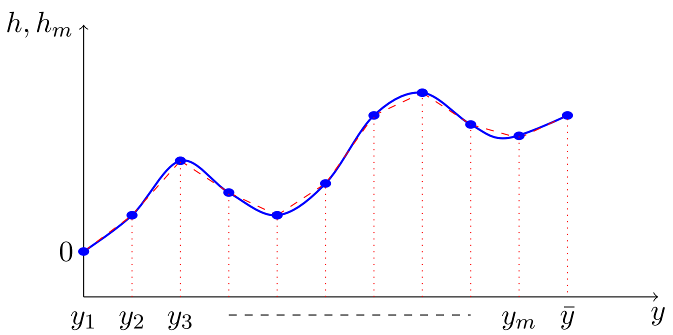


Figure 8: Approximation of function h (solid, blue) by piecewise linear function h_m (dashed, red).

The Gateaux differential is linear in the reform direction h and hence continuous. We therefore have

$$\lim_{m\to\infty} R_\tau(0,h_m) \quad = \quad R_\tau(0,h)\ , \tag{21}$$

i.e., the Gateaux differential for direction h_m converges to the Gateaux differential for direction h. We now provide a characterization of $\lim_{m\to\infty} R_\tau(0,h_m)$. By the linearity of the Gateaux differential, we have

$$R_\tau(0,h_m) \quad = \quad \textstyle\sum_{k=1}^{m} \tau_k R_\tau(0,h_k^s) = \sum_{k=1}^{m} \tau_k R_\tau^s(0,\ell_k,y_k)\ .$$

For m large and $\ell_k = \frac{\bar{y}}{m}$ close to zero, a first-order Taylor approximation moreover gives

$$R_\tau^s(0,\ell_k,y_k) \approx \ell_k\ R_{\tau\ell}^s(0,0,y_k) = \ell_k\ \mathcal{R}(y_k)\ .$$

The approximation is perfect in the limit case $m\to\infty$ or, equivalently, $\ell_k = \frac{\bar{y}}{m} \to 0$. Therefore,

$$\begin{aligned}
\lim_{m\to\infty} R_\tau(0,h_m) \quad &= \quad \lim_{m\to\infty} \sum_{k=1}^{m} \tau_k \ell_k \mathcal{R}(y_k) \\
&= \quad \lim_{m\to\infty} \sum_{k=1}^{m} [y_{k+1}-y_k]\ \ \tau_k\ \mathcal{R}(y_k) \\
&= \quad \int_{y\in\mathcal{Y}} h'(y)\ \mathcal{R}(y)\ dy\ ,
\end{aligned}$$

where the last term is the Riemann integral that gives the marginal revenue effect of a reform in direction h, and $h' : y \mapsto h'(y)$ is the change of the marginal tax rate at income y due to a unit increase in τ. To see this, note first that the term in the second line is the limit of a Riemann sum, the latter involving the step function $y \mapsto \tau_k\ \mathcal{R}(y_k)$ for $y \in [y_k, y_k+\ell_k]$. Second note that both h and h_m are continuous functions on a compact interval. Hence, they have a bounded variation and are therefore differentiable almost everywhere. Moreover, we have $\lim_{m\to\infty} \ell_k = 0$ and therefore, for any y in the interior of bracket k,

$$\tau_k = h_m'(y) \quad \longrightarrow_{m\to\infty} \quad h'(y)\ .$$

This completes the derivation of Equation (9) which appears in Proposition 2.

We now show that the conclusion of Theorem 2 extends to all continuous reform directions. We first note that (21) implies that for any $\varepsilon > 0$, there is an $\tilde{m}(\varepsilon) \in \mathbb{R}$ such that, for $m > \tilde{m}(\varepsilon)$

$$R_\tau(0, h) - R_\tau(0, h_m) \quad < \quad \varepsilon \; . \tag{22}$$

To complete the proof of Corollary 1, we proceed by contradiction. Suppose that the conditions of Theorem 2 apply and that there is a continuous reform direction h that is Pareto-improving, i.e., that satisfies

$$R_\tau(0, h) - \max_{y \in \mathcal{Y}} \; h(y) \quad = \quad 2\varepsilon \; ,$$

for some $\varepsilon > 0$. Then, by (20) and (22), for $m > \max\{\hat{m}(\varepsilon), \tilde{m}(\varepsilon)\}$, there is also an m-bracket reform such that $R_\tau(0, h_m) - \max_{y \in \mathcal{Y}} \; h_m(y) > 0$, i.e., that is Pareto-improving. But this is impossible by Theorem 2. The assumption that there is a Pareto-improving direction h in the class of continuous functions on $[0, \bar{y}]$ has therefore led to a contradiction and must be false.

B The revenue-maximizing tax cut

Fix a status quo tax-transfer system (T_0, c_0) and a corresponding revenue function $\mathcal{R}(y)$. By Proposition 2, a reform direction h is Pareto-improving if

$$\Pi(h) \quad := \quad \int_0^{\bar{y}} h'(y)\mathcal{R}(y)dy - \max_{y \in y_0(\Theta)} \int_0^y h'(z)dz \; , \tag{23}$$

is strictly positive. Then, a small reform in direction h provides a net revenue gain in excess to what is needed to compensate the agents facing the largest tax increase. In the following, we solve for the reform direction h_R that maximizes Π over the set of functions such that $h' : [0, \bar{y}] \to [-a, a]$ for some fixed $a > 0$. To simplify the exposition, we impose further assumptions that are satisfied in the context of our application in Section 5 (see Figure 7).

Assumption 1 *Let $y_0(\Theta) = [0, \bar{y}]$. There is a unique pair of income levels (y_s, y_t) with $0 < y_s < y_t < \bar{y}$ such that $\mathcal{R}(y)$ is strictly decreasing on $(0, y_s)$ and on $(y_t, \bar{y})$. Let $y_m = \frac{y_s + y_t}{2}$. There is a number $r \in (0, 1)$ so that:*

(i) $\mathcal{R}(y_s) = \mathcal{R}(y_t) = r$,

(ii) $\mathcal{R}(y) \in (r, 1)$ *for each* $y \in (0, y_s)$ *and for each* $y \in (y_m, y_t)$,

(iii) $\mathcal{R}(y) < r$ *for each* $y \in (y_s, y_m)$,

(iv) $\mathcal{R}(y) \in [0, r)$ *for each* $y \in (y_t, \bar{y})$,

In words: The monotonicity condition on $y \mapsto \mathcal{R}(y)$, see Theorems 1 and 2, is satisfied for incomes lower than y_s and for incomes higher than y_t, but violated for incomes between y_s and y_t. By (i), at y_s and y_t, marginal tax rates also satisfy the boundedness conditions in those Theorems. By (ii), for incomes lower than y_s and for incomes in the designated phase-out range (y_m, y_t), marginal tax rates are not inefficiently low. By (iii), for incomes in the designated phase-in range, the revenue potential is bounded by the one at y_s: y_s is a maximum of $y \mapsto \mathcal{R}(y)$ over this range of incomes. Finally, by (iv), for incomes larger than y_t, marginal tax rates are not inefficiently high and, over this income range, $y \mapsto \mathcal{R}(y)$ obtains a maximum at y_t.

Proposition 3 *Fix* $a > 0$. *Under Assumption 1, the revenue-maximizing tax cut* h_R *is, for any* $a > 0$, *given by a two-bracket tax cut with*

$$h'_R(y) = \begin{cases} 0 & \textit{for} \quad y \in [0, y_s) \ , \\ -a & \textit{for} \quad y \in (y_s, y_m) \ , \\ a & \textit{for} \quad y \in (y_m, y_t) \ , \\ 0 & \textit{for} \quad y \in (y_t, \bar{y}] \ . \end{cases} \tag{24}$$

Moreover, $\Pi(h_R) = aI(T_0, c_0)$ *with*

$$I(T_0, c_0) = \int_{y_m}^{y_t} \mathcal{R}(y)dy - \int_{y_s}^{y_m} \mathcal{R}(y)dy \ . \tag{25}$$

Proof To solve for the revenue-maximizing tax cut, we proceed in three steps. First, we show that $\max_{y \in y_0(\Theta)} h_R(y)$ equals zero. Second, we solve for the reform that maximizes $\Pi(h)$ subject to (i) $\max_{y \in y_0(\Theta)} h(y) = 0$, (ii) $h'(y) \in [-a, a]$ for all $y \in [y_s, y_t)$, and (iii) the additional restriction that $h'(y) = 0$ for all incomes below y_s and above y_t. Third, we show that the solution to this more restricted problem also solves the original maximization problem.

Step 1. The normalization that system $T(0) = 0$ for any tax system that we consider, also implies $h(0) = 0$ for any reform direction that we consider. Therefore $\max_{y \in y_0(\Theta)} h_R(y) \geq 0$. To show that $\max_{y \in y_0(\Theta)} h_R(y) = 0$, we provide a proof by contradiction. For this purpose, assume that there is some $\varphi > 0$ such that $\max_{y \in y_0(\Theta)} h_R(y) = \varphi$ and denote by y^* the lowest income level such that $h_R(y) = \varphi$. Then, $y^* > 0$ and there must be an income $y' \in (0, y^*)$ such that $h_R(y) > 0$ for all $y \in (y', y^*)$. Consider a perturbed reform h_ε such that $h'_\varepsilon(y) = h'_R(y) - \varepsilon$ for all incomes (y', y^*) and $h'_\varepsilon(y) = h'_R(y)$ for all other incomes. The net revenue gain from this perturbed reform is

$$\Pi(h_\varepsilon) = \int_0^{\bar{y}} h'_\varepsilon(y) \mathcal{R}(y) dy - \varphi + \varepsilon(y^* - y') .$$

The derivative of $\Pi(h_\varepsilon)$ with respect to ε is

$$\frac{d\Pi(h_\varepsilon)}{d\varepsilon} |_{\varepsilon=0} = - \int_{y'}^{y^*} \mathcal{R}(y) dy + (y^* - y') > 0 ,$$

where the positive sign follows because $\mathcal{R}(y) < 1$ for any $y \in (0, \bar{y})$ by Assumption 1. This contradicts that assumption that Π obtains a maximum at h_R.

Step 2. Consider the problem to maximize $\Pi(h)$ over the set of functions h such that (i) $h'(y) \in [-a, a]$ for all $y \in (y_s, y_t)$, (ii) $h'(y) = 0$ for all $y \leq y_s$ and all $y \geq y_t$ and (iii) $h(y) = \int_0^y h'(z) dz = \int_{y_s}^y h'(z) dz \leq 0$ for all $y \in (y_s, y_t)$. Note that the function given in (24) satisfies these constraints.

We now consider a Lagrangian for a more relaxed problem that takes only the constraint $h(y_t) \leq 0$ into account. We argue below that a solution to this relaxed problem satisfies (i)-(iii).

$$\mathcal{L}(t) = \int_{y_s}^{y_t} h'(y) \mathcal{R}(y) dy - \mu \int_{y_s}^{y_t} h'(y) dy ,$$

where μ is a Lagrange multiplier. The solution to this restricted problem is given by a function $\tilde{h} : (y_s, y_t) \to [-a, a]$ and a value $\tilde{\mu}$ of the multiplier. For any $y \in (y_s, y_t)$, the derivative of $\mathcal{L}$ with respect to $h'(y)$ is given by

$$\frac{\partial \mathcal{L}}{\partial h'(y)} = \mathcal{R}(y) - \tilde{\mu} .$$

As the Lagrangian in linear in each $h'(y)$, the solution involves $\tilde{h}'(y)$ equal to the lower bound $-a$ for all y such that $\mathcal{R}(y) < \tilde{\mu}$, and $\tilde{h}'(y)$ equal to the upper bound a for all y such that $\mathcal{R}(y) > \tilde{\mu}$. Under Assumption 1, this is only consistent with $\tilde{h}(y_t) = \int_{y_s}^{y_t} \tilde{h}'(y)dy = 0$ if $\tilde{\mu} = r = \mathcal{R}(y_s)$. Then, $\tilde{h}'(y) = -a$ for all $y \in (y_s, y_m)$ and $\tilde{h}'(y) = a$ for all $y \in (y_m, y_t)$. Hence, $\tilde{h}$ equals the function given in (24), so it satisfies (i)-(iii). Consequently,

$$\Pi(\tilde{h}) = \int_{y_s}^{y_m} -a\mathcal{R}(y)dy + \int_{y_m}^{y_t} a\mathcal{R}(y)dy = aI(T_0, c_0) \ ,$$

with $I(T_0, c_0)$ given in (25). We also note that, as $\tilde{\mu}$ is strictly positive, the constraint $\tilde{h}(y_t) \leq 0$ is binding.

Step 3. It remains to show that we cannot increase Π further by allowing $h'(y) \in \{-a, a\}$ for incomes below y_s and above y_t, while respecting the constraint $\int_0^y h(z)dz \leq 0$ for all $y \in [0, \bar{y}]$. A repeated application of the arguments in Step 2, once for incomes below y_s, and once for incomes above y_t, exploiting the monotonicity of $y \mapsto \mathcal{R}(y)$ over these income ranges, shows that any candidate solution to this problem will take the form

$$h_R'(y) = \begin{cases} 0 & \text{for} \quad y < y_\alpha \ , \\ -a & \text{for} \quad y \in (y_\alpha, y_m) \ , \\ a & \text{for} \quad y \in (y_m, y_\beta) \ , \\ 0 & \text{for} \quad y > y_\beta \ . \end{cases}$$

where $y_\alpha \leq y_s$ and $y_\beta \geq y_t$. The constraint $h(\bar{y}) = \int_0^{\bar{y}} h_R(y)dy \leq 0$ is only satisfied if $y_\beta - y_m \leq y_m - y_\alpha$. Finally, choosing y_α and y_β to maximize $\Pi(h_R)$ subject to $y_\alpha \leq y_s$ and $y_\beta \geq y_t$ and $y_\beta - y_m \leq y_m - y_\alpha$, shows that $y_\alpha = y_s$ and $y_\beta = y_t$ is an optimal choice.