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

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

June 4, 2020

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, Andreas Peichl, Dominik Sachs, Florian Scheuer, Nick Werquin and seminar audiences in Cologne, Munich, Paris, 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 and the Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) 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, 5, avenue Henry Le Chatelier, 91128 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 the 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 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 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 reform at all.

These findings are derived from a generic static model of taxation: Individuals derive utility from consumption and the generation of income requires costly effort. Individuals

[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 (2019).

[2]See Piketty (1997), Saez (2001), Golosov, Tsyvinski and Werquin (2014), or Jacquet and Lehmann (2016).

face a budget constraint that is shaped by a non-linear tax system. A predetermined tax system is in place and we consider reforms of that 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 tool that provides an answer. The tool is a test function which gives, for each income level y, the revenue implications 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 test function is bounded from below by 0 and from above by 1.

2. There is no Pareto-improving two-bracket reform if and only if the test function is non-decreasing.

Thus, all that is needed to test for Pareto efficiency is a sufficient statistics formula for the revenue implications of a small tax reform involving a modification of the marginal tax rates 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.

The introduction of the EITC. We look at the introduction of the EITC in the US through the lens of our framework. The introduction of the EITC in the mid-1970s 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. We then use our test function to investigate whether or not past reforms of the EITC in the US went into a Pareto-improving direction. We also check whether the reforms led to a Pareto-efficient tax system.

We find that, prior to the introduction of the EITC, the test function 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, leaving room for further Pareto-improvements by means of two-bracket reforms. 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.[3]

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.[4] 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, the EITC can be justified as being part of a policy that is, in a utilitarian sense, optimal.[5]

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.[6] The status quo plays no role in the theory of optimal taxation: what is optimal does not depend on what is currently in place. 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

[3]The inefficient pattern that existed in the mid-1970s has become less pronounced over time, but can still be detected in the US tax and transfer system of 2018, see Appendix E.

[4]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 low or middle income than to people with no income; see Stiglitz (1982), Choné and Laroque (2010), or Brett and Weymark (2017).

[5]Follow-up papers are Jacquet et al. (2013) or Hansen (2019). Saez and Stantcheva (2016) enrich the traditional welfarist approach to account for existing tax policy debates while maintaining the desirability of Pareto efficiency.

[6]See Piketty (1997), Saez (2001), Golosov et al. (2014), or Jacquet and Lehmann (2016). For earlier contributions to the analysis of tax reforms, see Feldstein (1976), Weymark (1981) and the review in Guesnerie (1995).

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.[7]

Outline. The remainder is organized as follows. The next section discusses the related literature. Section 3 contains our theory of Pareto-improving tax reforms, and the sufficient statistics formula for a model with labor supply responses at the intensive and the extensive margin. Section 4 contains the application of these formulas to the introduction of the EITC in the US. Formal proofs are relegated to the Appendix. The Online-Appendix contains sufficient statistics formulas for the welfare-implications of two-bracket reforms that are not Pareto-improving, and a political economy analysis of two-bracket reforms.

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).[8] Bierbrauer, Boyer and Peichl (2020) 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.[9] Failures of Pareto efficiency are also identified by the literature on the *inverse tax problem*.[10] This literature identifies the welfare function that is maximized by an observed tax schedule. If this approach yields a welfare function with

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

[8]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).

[9]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).

[10]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).

negative weights for some individuals, this indicates that the tax policy under study is incompatible with the maximization of a Paretian social welfare function.[11]

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. (2020), on the one hand, and the conditions by Werning (2007) and Lorenz and Sachs (2016), on the other, imply Pareto efficiency.[12]

We moreover derive sufficient statistics that can be used to check whether the necessary and sufficient 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 (2018) for reviews of this approach.

Our analysis of the introduction of the EITC in the US draws on the literature that provides estimates of the behavioral responses involved.[13] Bastian (2020) estimates the strength of responses to the 1975 EITC introduction, the reform that we focus on as well. There is a range of estimates and some of the conventional wisdom in the literature has recently been challenged.[14] We do not have to take a stance here. Our conclusion that the introduction of the EITC targeted an inefficiency in the US tax and transfer system is valid for all levels of the intensive-margin and extensive-margin labor supply elasticities that are discussed in the literature.

Our test for Pareto efficiency employs sufficient statistics for the revenue implications of a small tax reform. Pareto efficiency fails if such a reform is self-financing. A related discussion of the extent to which past reforms of the EITC have been self-financing can be found in Bastian and Jones (2019).

[11] In part C of the Appendix, we explain in detail what the conditions for Pareto efficiency in Theorems 1 and 2 imply for the inverse tax problem.

[12] 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 (see Section 3).

[13] 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).

[14] 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 and Jones (2019) and Kleven (2019) find smaller participation elasticities.

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 $\theta \in \Theta \subset \mathbb{R}^n$. 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 and increasing in the first argument, with partial derivative denoted by u_c. It is decreasing in the second argument. The cross-section distribution of θ is assumed to be atomless and represented by a cumulative distribution function F.

We keep the analysis general in this section but note that 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. The tax function T_0 assigns a tax payment to every earnings level. Without loss of generality, we let $T_0(0) = 0$.[15] We assume that T_0 is continuous. Otherwise, it can be an arbitrary non-linear tax function, possibly with kinks. Before the

[15] 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. In the literature, $T_0(y)$ is often referred to as a participation tax. It gives the tax that a person with income y has to pay, relative to a person who stays out of the labor market and has zero income, see, e.g., Kleven (2014).

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.

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. 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 assume that the intercept absorbs the reform-induced changes in tax revenue that we denote by $R(\tau, h)$. Thus,

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

The change in tax revenue is an endogenous object that depends on the behavioral responses to taxation. 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.[16] Thus, $R(\tau, h)$ 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.

We denote by $v(R(\tau, h), \tau, h, \theta)$ the indirect utility that a type θ individual realizes after a tax reform. We can use the analysis of *"Envelope theorems for arbitrary 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 and suppose that the problem in (1) has a unique solution. Then, by Corollary 4 in Milgrom and Segal (2002),

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

where the marginal consumption utility of type θ, $u_c(\cdot, \theta)$, is evaluated at $(C_1(y^*(\cdot)), y^*(\cdot))$, and $R_\tau(\tau, h)$ is the marginal effect of a change of the reform intensity τ on tax revenue.[17]

[16] 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.

[17] For a type θ so 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.

More formally, it is the Gateaux differential of tax revenue in direction h.[18] The envelope theorem covers, in particular, 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. Suppose, for concreteness, that a small reform step has a positive impact on tax revenue, $R_\tau(\tau, h) > 0$. A taxpayer benefits from the reform if and only if this revenue gain outweighs the additional tax payment $h(y^*(\cdot))$. Hence, a small increase of τ is Pareto-improving if and only if

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

where, for given τ and h, $y^*(\Theta)$ is the image of the function y^*. Analogously, lowering τ is Pareto-improving if and only if

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

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

$$R_\tau(\tau, h) - \max_{y \in y^*(\Theta)} h(y) \leq 0 \,, \tag{6}$$

and

$$R_\tau(\tau, h) - \min_{y \in y^*(\Theta)} h(y) \geq 0 \,. \tag{7}$$

3.2 The main results

Is it possible to raise everyone's utility by increasing or lowering the marginal tax rates in a finite number of income brackets? Theorems 1 and 2 provide answers to this question. Before we can state these results, we need some additional notation.

We describe a multi-bracket reform as a way of combining a collection of m single-bracket reforms, $\{(\tau_k, \ell_k, y_k)\}_{k=1}^m$, where a generic single-bracket reform (τ_k, ℓ_k, y_k) is characterized by the income level y_k at which the bracket starts, the length of the bracket ℓ_k and the change of marginal tax rates for incomes in the bracket, τ_k. Thus, if

[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 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 $R_\tau(0, h) := \partial \mathcal{R}(T_0, h)$ and $R_\tau(\tau', h) := \partial \mathcal{R}(T_0 + \tau' h, h)$.

$y \in (y_k, y_k + \ell_k)$, then $T_1'(y) = T_0'(y) + \tau_k$. We will trace the welfare implications of multi-bracket reforms back to the properties of single-bracket reforms.

Note that a single-bracket reform can be equivalently described as a pair (τ_k, h_k), where the function h_k is such that

$$h_k(y) = \begin{cases} 0, & \text{if} \quad y < y_k \ , \\ y - y_k, & \text{if} \quad y \in [y_k, y_k + \ell_k] \ , \\ \ell_k, & \text{if} \quad y > y_k + \ell_k \ . \end{cases}$$

It will prove convenient to have separate notation for the revenue implications of single-bracket reforms. For such reforms, we write $R^s(\tau_k, \ell_k, y_k)$ rather than $R(\tau_k, h_k)$. 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_k and ℓ_k close to zero,

$$R^s_{\tau\ell}(0, 0, y_k) \ \tau_k \ \ell_k$$

is a good approximation of $R^s(\tau_k, \ell_k, y_k)$, i.e., of the reform's revenue implications. Thus, the cross-derivative can be interpreted as a measure of how much revenue can be raised by a small single-bracket reform.

Our analysis uses the fact that, with an atomless type distribution, $R^s_\tau(0, 0, y_k) = 0$ for all y_k: A reform that slightly increases marginal taxes on an interval of length zero affects tax liabilities for a zero mass of incomes and, hence, does not generate tax revenue.

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

Theorem 1 states necessary conditions for the Pareto efficiency of a tax system. These conditions involve $y \mapsto R^s_{\tau\ell}(0, 0, y)$, which gives the revenue implications of a small single-bracket reform as a function of the income level at which marginal tax rates are changed.

The first condition is that $R^s_{\tau\ell}(0, 0, y) \geq 0$ for all y. Hence, a reform involving an increase of marginal tax rates, $\tau_k > 0$, 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 $R^s_{\tau\ell}(0, 0, y) < 1$ 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 raising τ_k 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 third condition is that the function $y \mapsto R^s_{\tau\ell}(0, 0, y)$ is non-increasing. It is derived from an analysis of reforms that involve two brackets:

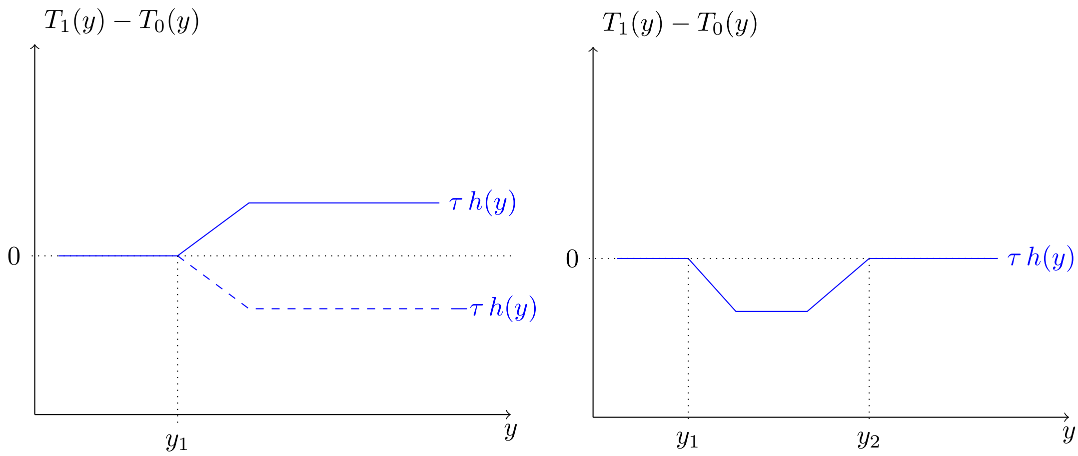
$$h_j(y) = \tau_1 \ h_1(y) + \tau_2 \ h_2(y) \ , \tag{8}$$

$$h_1(y) = \begin{cases} 0, & \text{if} \quad y \leq y_1 \ , \\ y - y_1, & \text{if} \quad y \in (y_1, y_1 + \ell_j \, \ell_1) \ , \\ \ell_j \, \ell_1, & \text{if} \quad y \geq y_{a1} + \ell_j \, \ell_1 \ . \end{cases}$$

and

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

Thus, a two-bracket reform links two single-bracket reforms (τ_1, ℓ_1, y_1) and (τ_2, ℓ_2, y_2) in a particular way: marginal tax rates change by $\tau_j \, \tau_1$ for incomes in the first bracket and by $\tau_j \, \tau_2$ for incomes in the second bracket. We call a two-bracket reform symmetric if $\tau_1 \ell_1 + \tau_2 \ell_2 = 0$. A symmetric two-bracket reform with a phase-in range where marginal taxes are reduced and a subsequent phase-out where marginal taxes are increased has $\tau_1 < 0$ and $\tau_2 > 0$. We refer to such reforms as *symmetric two-bracket tax cuts.* This choice of terminology reflects that the reform does not increase anyone's tax burden and that all people with an income between the endpoints of the two brackets get a tax cut, see Figure 1 for an illustration.



(a) Reforms with one bracket.

(b) A symmetric two-bracket tax cut.

Figure 1: Income tax reforms with one bracket and two brackets.

Our construction of two bracket reform facilitates an analysis of the limit case $\tau_j \to 0$ and $\ell_j \to 0$, see Figure 2 for an illustration. As τ_j 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_j is send to zero, while the ratio of their lengths is kept constant at $\frac{\ell_1}{\ell_2}$.

Proposition 1 *The following statements are equivalent:*

1. *The function $y \mapsto R^s_{\tau\ell}(0, 0, y)$ is increasing over some range.*

2. *A Pareto improvement can be realized with a symmetric two-bracket tax cut.*

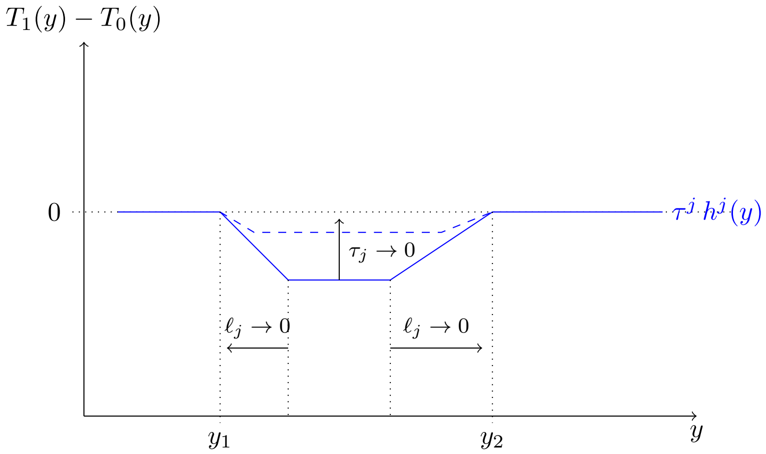


Figure 2: A small symmetric two-bracket reforms with a tax cut.

In light of this Proposition, 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 symmetric two-bracket tax cuts. As we show formally in the proof of Proposition 1, if $y \mapsto R^s_{\tau\ell}(0,0,y)$ is increasing, such a tax cut 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 $y \mapsto R^s_{\tau\ell}(0,0,y)$ must be non-increasing is an analogue to the condition $R^s_{\tau\ell}(0,0,y) \geq 0$, for all y. The latter rules out the existence of self-financing tax cuts for single-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 these findings, 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 $y \mapsto R^s_{\tau\ell}(0,0,y)$ is non-increasing, bounded from below by 0 and bounded from above by 1, then there is no Pareto-improving direction in the class of reforms with finitely many brackets.*

According to Theorem 2, if there is no Pareto-improving reform with one or two brackets, then there is no Pareto-improving direction with any number m of brackets. Put differently, the tax system admits a Pareto-improving reform with finitely many brackets only if there also is a Pareto-improving reform involving either one or two brackets. Thus, ruling out Pareto-improving one-bracket reforms and Pareto-improving two bracket reforms guarantees Pareto efficiency.

The formal proof consists in showing that, when the conditions of Theorem 2 are fulfilled, any reform direction h with finitely many brackets is such that

$$R_\tau(\tau, h) - \max_{y \in y^*(\Theta)} h(y) \leq 0 \tag{9}$$

and

$$R_\tau(\tau, h) - \min_{y \in y^*(\Theta)} h(y) \geq 0 \,, \tag{10}$$

implying that no Pareto-improving reform direction can be found. The proof exploits the linearity of the Gateaux differential, but otherwise uses only elementary arguments.

The difficulty in the proof comes from the analysis of m-bracket reforms with a set of brackets where marginal tax rates are raised and another set of brackets where marginal tax rates are lowered.[19] Moreover, any such case comes with various possibilities as to how peoples' total tax payments can be affected. For instance, a symmetric two-bracket reform with a tax cut has a first bracket with lower marginal tax rates and a second bracket with increased rates, chosen in such a way that no ones' total tax payment goes up. There are other two-bracket reforms so that the total tax payment goes up for some and goes down for others. There are also two-bracket reforms where the total tax payment goes up for everyone. Obviously, when we consider m rather than two brackets, the number of possible constellations is much larger. Proving Theorem 2 amounts to showing that none of these many constellations can be Pareto-improving. Our proof of Theorem 2 achieves this purpose in a concise way, but otherwise does not convey much economic intuition.

A broad intuition is that Pareto-improving reforms have to be self-financing. If $R^s_{\tau\ell}(0, 0, y)$ is positive, a reform that lowers the tax burden for some people can only be self-financing if it involves an increase of the tax burden for others. If $y \mapsto R^s_{\tau\ell}(0, 0, y)$ is non-increasing, then, to every person who benefits from tax cut, there has to be a richer person whose tax burden increases by an amount that exceeds the initial tax cut. But this makes it impossible to have this richer person included in the set of reform beneficiaries; i.e., budget balance cannot be achieved in a Pareto-improving way.

3.3 Discussion

Theorems 1 and 2 and Proposition 1 provide a characterization of Pareto-improving reforms and of Pareto-efficient tax systems. By Theorem 1, if a tax system is Pareto

[19] An m-bracket reform so that marginal tax rates are lowered in all m brackets has an effect on revenue and on the taxpayers' indirect utilities which is simply a convex combination of the effects that any such reform would have on a stand-alone basis. If none of the reforms in the package is Pareto-improving, the whole package cannot be Pareto-improving. The same is true for an m-bracket reform so that marginal tax rates are increased in all m brackets.

efficient, it must be that $0 \leq R^s_{\tau\ell}(0,0,y) \leq 1$ for all y. Otherwise, there exists a Pareto-improving one-bracket reform. By Theorem 1 and Proposition 1, if a tax system is Pareto-efficient, the function $y \mapsto R^s_{\tau\ell}(0,0,y)$ must be non-decreasing. Otherwise, a Pareto improvement can be realized by means of a symmetric two-bracket tax cut. These findings imply, in particular, that a tax system may be inefficient even though there is no way to Pareto-improve by means of a simple tax reform that involves only one bracket. Thus, there are Pareto improvements that can only be realized with two-bracket reforms. Does the consideration of three, four or five brackets enlarge the scope for Pareto improvements even further? By Theorem 2, the answer is "no." If there is neither a Pareto-improving one-bracket reform nor a Pareto-improving two-bracket reform, then there is no Pareto-improving reform direction at all.

Making use of this characterization. These insights can be used in combination with sufficient statistics for the test function $y \mapsto \mathcal{R}^s_{\tau\ell}(0,0,y)$. Given a tax system that is to be analyzed, this approach allows to provide answers to a variety of questions: Is the function $y \mapsto \mathcal{R}^s_{\tau\ell}(0,0,y)$ throughout between zero and one or, equivalently, is there a Pareto-improving one-bracket reform? Is the function $y \mapsto \mathcal{R}^s_{\tau\ell}(0,0,y)$ non-increasing or, equivalently, is there a Pareto-improving two-bracket reform? One can also study the function $y \mapsto \mathcal{R}^s_{\tau\ell}(0,0,y)$ at different dates, in order to see whether a reform alleviated or aggravated an inefficiency or, possibly, created a new one.

Different models of taxation give rise to different test functions. Thus, the concrete specification of $y \mapsto \mathcal{R}^s_{\tau\ell}(0,0,y)$ will depend on the application of interest and on a choice of what model to use for this application. This important point may have been buried by the general analysis that led to Theorems 1 and 2 and Proposition 1. We illustrate it with two examples.[20]

First, in the model of Diamond (1998) with $u(c,y,\theta) = c - \left(\frac{y}{\theta}\right)^{1+\frac{1}{\varepsilon}}$ and $\theta \in \Theta \subset \mathbb{R}_+$,

$$R^s_{\tau\ell}(0,0,y) = 1 - F(\theta_0(y)) - \frac{T_0'(y)}{1-T_0'(y)} \left(1 + \frac{1}{\varepsilon}\right)^{-1} f(\theta_0(y))\,\theta_0(y)\;, \tag{11}$$

where θ_0 is the inverse of the function y_0. Thus, $\theta_0(y)$ is the type who earns y in the status quo. In applications, one usually observes the distribution of incomes, and not the distributions of types. Therefore, applications are often based on an alternative formulation of equation (11) that refers to the distribution of earnings, $F_y(y) := F(\theta_0(y))$,

$$R^s_{\tau\ell}(0,0,y) = 1 - F_y(y) - \frac{T_0'(y)}{1-T_0'(y)}\,\varepsilon\, y f_y(y)\;\;. \tag{12}$$

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. We present such a

[20] A derivation of the following equations (11) and (12) can be found in Appendix D.2.1. Specifically, both are Corollaries to Proposition D.1 in Appendix D.

framework and derive a sufficient-statistics formula characterizing $y \mapsto R^s_{\tau\ell}(0,0,y)$ in the Appendix (see Proposition D.1 in Appendix C). 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. Here, for ease of exposition, we focus again on the case of quasi-linear preferences and iso-elastic effort costs. This is also the specification that we will use in our analysis of the EITC in the subsequent section. Hence, suppose that

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

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,

$$R^s_{\tau l}(0,0,y) = 1 - F_y(y) - \varepsilon\, y\, f_y(y)\, \frac{T_0'(y)}{1-T_0'(y)} - \int_y^\infty f_y(y')\, \pi_0(y')\, \frac{T_0(y')}{y'-T_0(y')}\, dy' \; , \quad (13)$$

where $\pi_0(y)$ is an extensive-margin (participation) elasticity.[21] 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.

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, $R^s_{\tau\ell}(0,0,y)$ is given by equation (11). If the expressions on the right hand side of (11) are taken to be differentiable in y, then the condition that $y \mapsto R^s_{\tau\ell}(0,0,y)$ 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 $\frac{f(\theta_0(y))}{1-F(\theta_0(y))}$. Both Werning (2007) and Lorenz and Sachs (2016) present such equations. Thus, their findings are implied by our result in Theorem 1 that monotonicity of the function $y \mapsto R^s_{\tau\ell}(0,0,y)$ is necessary for Pareto efficiency.

Tagging. Our analysis can be extended to allow for tagging.[22] 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, earnings subsidies for lone mothers may be different from those for childless individuals. The above analysis of Pareto-efficient taxation can then be applied separately for each group. Any Pareto-improving reform is self-financing within the relevant group. Thus, they can be analyzed without having to worry about distributive consequences across groups.

[21] For a formal derivation of equation (13), see Appendix D.2.2.

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

Reforms with winners and losers. If a reform is not Pareto-improving, then there are individuals who are made better off and others who are made worse off. This raises the question whether the reform is worthwhile from a welfare perspective, i.e., whether the gains of the winners exceed the costs of the losers. From a political economy perspective, the question is whether the winners form a majority. We provide answers to these questions in Appendix C, where we characterize the welfare and the political economy implications of reforms.

4 Application: The introduction of the EITC

We now relate our insights on Pareto-improving tax reforms to the introduction and subsequent expansion of the EITC in the US in the mid 1970s. More specifically, we will use Theorems 1 and 2 in combinations with the sufficient statics formula in (13) to see whether the introduction of the EITC was a reform in a Pareto-improving direction.

4.1 Background

The introduction of the EITC in 1975 was a response to a "poverty trap":[23] In the 1960s, new welfare programs had been introduced as part of President Johnson's "War on poverty." The new programs provided more generous benefits to families with low incomes. The benefits were phased out with income in a way that implied high effective marginal tax rates for many low-income families, exceeding 80% in many cases (see Figure 3 below). In the following decade, the share of welfare recipients increased substantially. By the early 1970s, finding ways to reduce the "poverty trap" and to increase work incentives was considered a pressing concern.

The US Congress enacted the *Earned Income Tax Credit (EITC)* as a temporary policy for the year 1975.[24] As described in Bastian (2020), this was a substantial policy change that affected a large share of the population.[25] 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 working taxpayers with dependent children. Later, the EITC became a permanent policy. Over the following decades, there were several expansions.[26]

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

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

[25] According to CPS data, about 47% of single parents and 42% of childless singles had earned incomes in the EITC range (i.e., strictly positive and below $8,000$ USD).

[26] For example, more generous credits for parents with two and three or more children were introduced.

Data description. 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.

We use data on the tax and transfer system and various welfare programs for the years 1974 and 1975 (see Table B.1 in Appendix B.1 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 called Food Stamps). The details of AFDC varied across states, so that a unified treatment for the US at large is not possible. We instead 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. Hence, we focus on the largest subgroup of single parents: those with two children.[27] Figure 3 shows effective marginal tax rates and average tax rates, $y \mapsto \frac{T_0(y)}{y}$, for single parents (left panel) and for childless singles (right panel) before the reform in 1974. At low incomes, both marginal tax rates and average tax rates were higher for single parents. The reason is that the phasing-out of AFDC and SNAP transfers implied an income range with exceptionally high marginal tax rates. The dotted vertical lines in both panels of Figure 3 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).

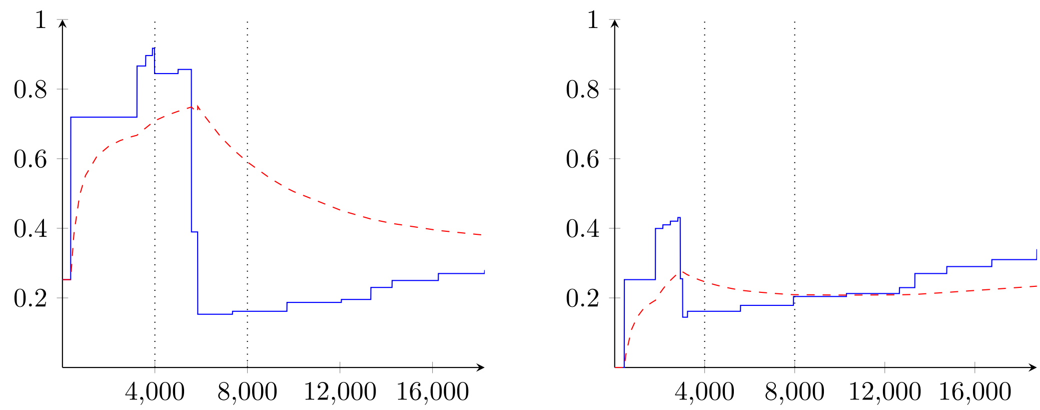
We estimate the 1974 and 1975 income distributions based on data from the March 1975 and March 1976 Current Population Survey (CPS), respectively. For this purpose, we consider the sample of non-married individuals aged 25 to 60 who do neither co-habit with an unmarried spouse nor with another adult family member. We partition this sample into childless singles and single parents.[28] 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.[29] For

In the 1990s, US authorities also extended the EITC program to childless workers and expanded the eligibility thresholds for married taxpayers. See Hoynes (2019) for a review.

[27]In our data, both the modus and the median number of children in single-parent households was two, and the arithmetic mean was close to two.

[28]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 ($N = 453$): Figure B.3 in Appendix B shows that the two distributions are essentially identical.

[29]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). Appendix B provides further details and shows that our qualitative results remain unchanged if we explicitly account for the assets and capital income tests.



(a) Single parents (b) Childless singles

Figure 3: US income tax and transfer schedules in 1974 for single parents and childless singles.

Notes: Figure 3 shows the 1974 effective marginal tax $T_0'(y)$ (blue lines) and average tax $\frac{T_0(y)}{y}$ (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 Appendix B.1 for details).

our benchmark analysis, we estimate the income distributions for both groups using a non-parametric kernel density estimation.[30]

We draw on a rich literature providing estimates of labor supply responses at the intensive and the extensive margin. There is a range of estimates that are supported by this literature (for a more detailed discussion, see Appendix B.1). As will become clear, our conclusions on the desirability of the EITC are robust, in the sense that they do not depend on how we select from this set. Initially, we follow Chetty, Guren, Manoli and Weber (2013) and present results for an intensive-margin elasticity of 0.33, both for single parents and childless singles. Extensive-margin elasticities are assumed to be decreasing with income in both groups. For single parents, this elasticity is taken to fall from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, giving rise to an average value of 0.52. For childless singles, we assume that it falls from 0.4 to 0.1, so that the average value is around 0.25. Below, we also present a sensitivity analysis that explores alternative assumptions.

4.2 Empirical results

We present the calibrations of two functions $y \mapsto \mathcal{R}^s_{sp}(y)$ and $y \mapsto \mathcal{R}^s_{cs}(y)$ giving, respectively, the revenue implications of small one-bracket reforms affecting *s*ingle *p*arents and

[30]Figure B.1 in Appendix B shows the density of this distribution. Results with gamma and lognormal distributions in Appendix B.2 show that this is not affecting our results.

childless singles.[31]

Figure 4 plots $y \mapsto \mathcal{R}^s_{sp}(y)$ and $y \mapsto \mathcal{R}^s_{cp}(y)$. It indicates that there were Pareto-improving reforms of the tax and transfer system applying to single parents. This is most pronounced for incomes around $4,000$ USD. For these incomes, $\mathcal{R}^s_{sp}(y) < 0$ indicates that a lowering of marginal tax rates would have been self-financing and Pareto-improving. For incomes above $4,000$ USD, $y \mapsto \mathcal{R}^s_{sp}(y)$ is increasing so that also a symmetric two-bracket tax cut would have been self-financing and Pareto-improving.

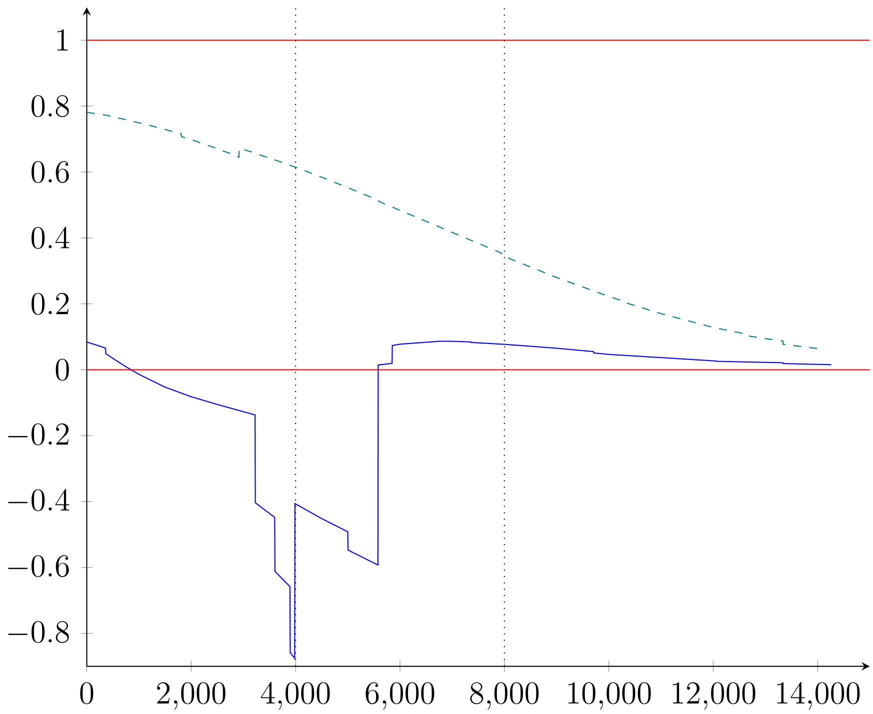


Figure 4: Pareto (in)efficiency of the 1974 US tax and transfer system.

Notes: Figure 4 shows the marginal revenue functions for single parents $\mathcal{R}^s_{sp}(y)$ (blue, solid) and childless singles $\mathcal{R}^s_{cs}(y)$ (teal, dashed) 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). Intensive-margin elasticities are fixed at 0.33, both for single parents and childless singles. Extensive-margin elasticities are assumed to be decreasing with income in both groups: for single parents, this elasticity is taken to falls from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52. For childless singles, it is assumed to fall from 0.4 to 0.1, with an average value around 0.25.
Source: Authors' calculations (see Appendix B.1 for details).

Interestingly, these findings do not extend to childless singles: throughout, we observe that $0 \leq \mathcal{R}^s_{cs}(y) \leq 1$, so that, with one-bracket reforms, there is no Pareto-improving direction. There is a small non-monotonicity for incomes around $3,000$ USD that could,

[31]Note that $\mathcal{R}^s(y)$ is a shorthand for the expression $R^s_{\tau l}(0, 0, y)$ introduced in the previous section. Henceforth, we use this more concise notation.

in principle, rationalize a two-bracket reform; but the scope is very limited.

Did the EITC introduction lead to a Pareto-efficient tax system? Our analysis of the (post-reform) 1975 tax and transfer system suggests that, for single parents, the answer is "no". Figure 5 illustrates the change in the marginal tax rates from 1974 to 1975, mainly reflecting the introduction of the EITC, with a phase-in range where marginal tax rates went down and a phase-out range where they went up. Figure 6 shows the post-reform calibration of $y \mapsto \mathcal{R}^s_{sp}(y)$. There is still a range where this function lies below zero and is increasing, indicating the scope for further Pareto-improving one- and two-bracket reforms. Interestingly, another two-bracket reform took place in 1979. It involved an expansion of the EITC for incomes between $4,000$ and $10,000$ USD. Figure B.5 in Appendix B.2 depicts the test function $y \mapsto \mathcal{R}^s_{sp}(y)$ for all years between the EITC introduction in 1975 and its first expansion in 1979. These figures shows that the inefficiency in the tax and transfer system was not eliminated by the reforms in the 1970s.

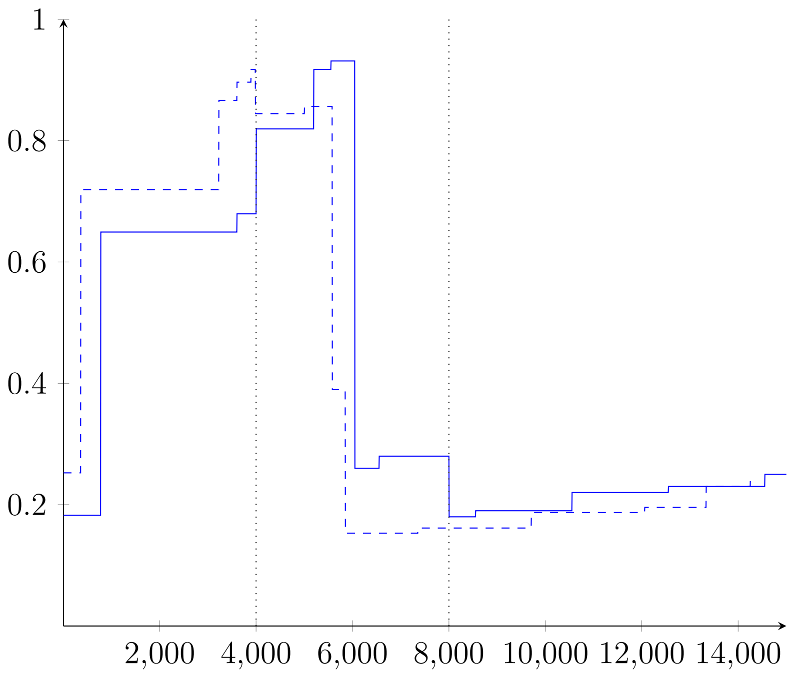


Figure 5: US income tax and transfer schedules in 1974 and 1975 for single parents.

Notes: Figure 5 shows the effective marginal tax rates for single parents before (dashed) and after (solid) the introduction of the EITC in 1975 as functions of earned income in 1974 USD and 1975 USD, respectively. 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 Appendix B.1 for details).

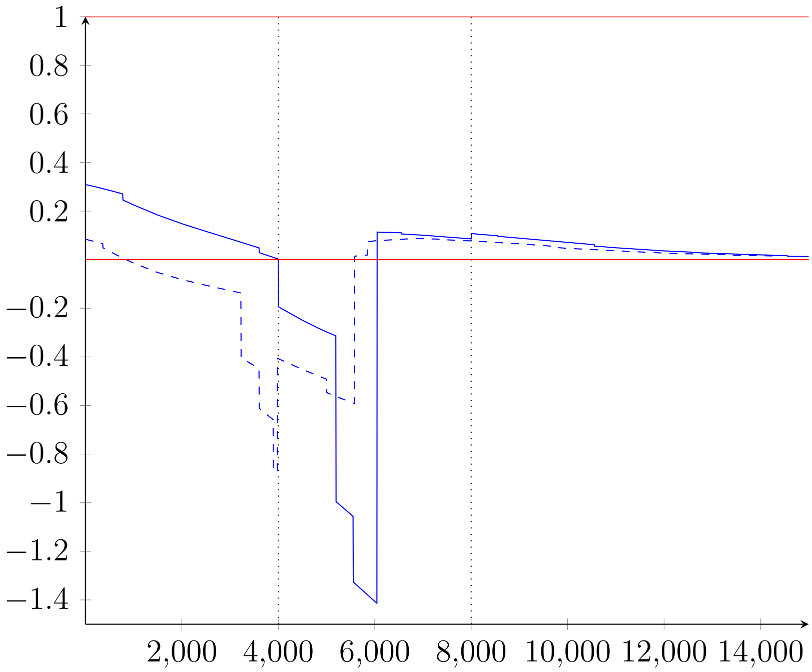


Figure 6: Pareto (in)efficiency of the 1975 US tax and transfer system for single parents.

Notes: Figure 6 shows the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$ before (dashed) and after (solid) the 1975 reform as functions of earned income in 1974 USD and 1975 USD, respectively. 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). Intensive-margin elasticities are fixed at 0.33, both for single parents and childless singles. Extensive-margin elasticities are assumed to be decreasing with income in both groups: for single parents, this elasticity is taken to falls from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52. For childless singles, it is assumed to fall from 0.4 to 0.1, with an average value around 0.25.
Source: Authors' calculations (see Appendix B.1 for details).

4.3 Sensitivity analysis

How sensitive are these conclusions to alternative assumptions on labor supply elasticities? To answer this question, Figures 7 and 8 plot the test function $y \mapsto \mathcal{R}^s_{sp}(y)$ for a variety of cases. The figures show the robustness of our finding that, as of 1975, there was scope for Pareto-improving tax cuts for incomes around $4,000$ USD.

More specifically, Figure 7 varies the intensive-margin elasticity ε from 0 to 0.5, while the participation elasticity is assumed to decrease from 0.6 at low incomes to 0.35 at high incomes, as in our benchmark analysis in Figure 4. We always detect a violation of Pareto-efficiency, even if we assume that there are literally no behavioral responses at the intensive margin, i.e., for $\varepsilon = 0$.

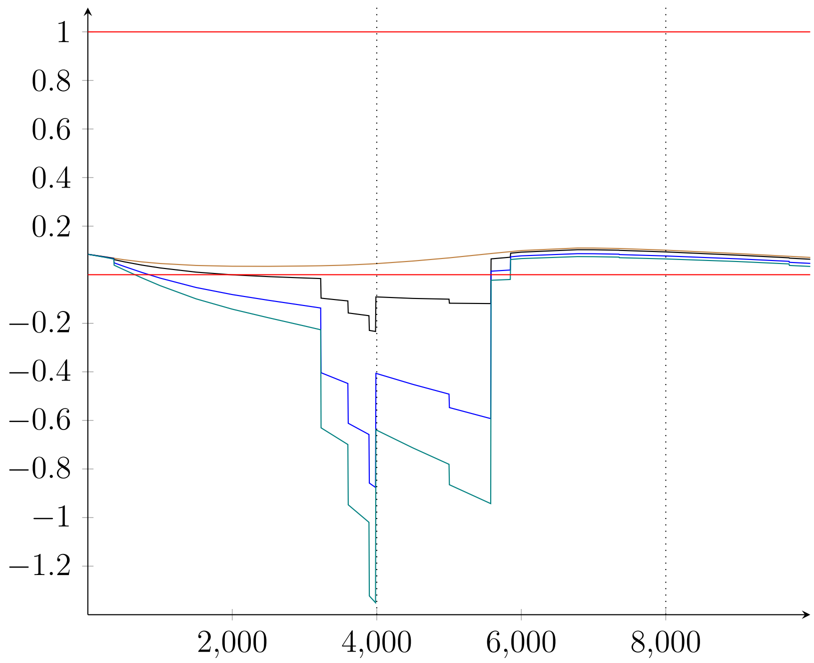


Figure 7: Pareto (in)efficiency of the 1974 US tax and transfer system, varying intensive-margin elasticities.

Notes: Figure 7 shows the marginal revenue function for single parents $y \mapsto \mathcal{R}^s_{sp}(y)$ before the 1975 reform as a function 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). As in Figure 4, extensive-margin elasticities decrease with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52. For childless singles, it is assumed to fall from 0.4 to 0.1, with an average value around 0.25. The intensive-margin elasticity varies from 0 (brown line) to 0.1 (black line), 0.33 (blue line) and 0.5 (teal line).
Source: Authors' calculations (see Appendix B.1 for details).

Figure 8 keeps the intensive elasticity constant at $\varepsilon = 0.33$ as in our benchmark analysis in Figure 4, and explores alternative assumptions about extensive margin responses, as captured by the elasticity π_0: a constant extensive margin elasticity of 0 or 0.4; one that is falling from 0.3 at low incomes to 0.1 at higher incomes; one that is falling from 0.6 to 0.35, one that is falling from 0.7 to 0.45 and one that is falling from 1 to 0.75. The takeaway is that in all cases, even with an extensive-margin elasticity of zero (brown line in Figure 8), we can identify Pareto-improving reforms with one and two brackets.

These results indicate that the existence of Pareto-improving reforms is robust with respect to the strength of behavioral responses to taxation. The introduction of the EITC did not fully remove these inefficiencies, leaving room for further Pareto improvements by means of two-bracket reforms. In Appendix E, we also present an analysis of the 2018 tax and transfer system. Our analysis suggests that the inefficiencies have been mitigated over time, but have not been fully eliminated.

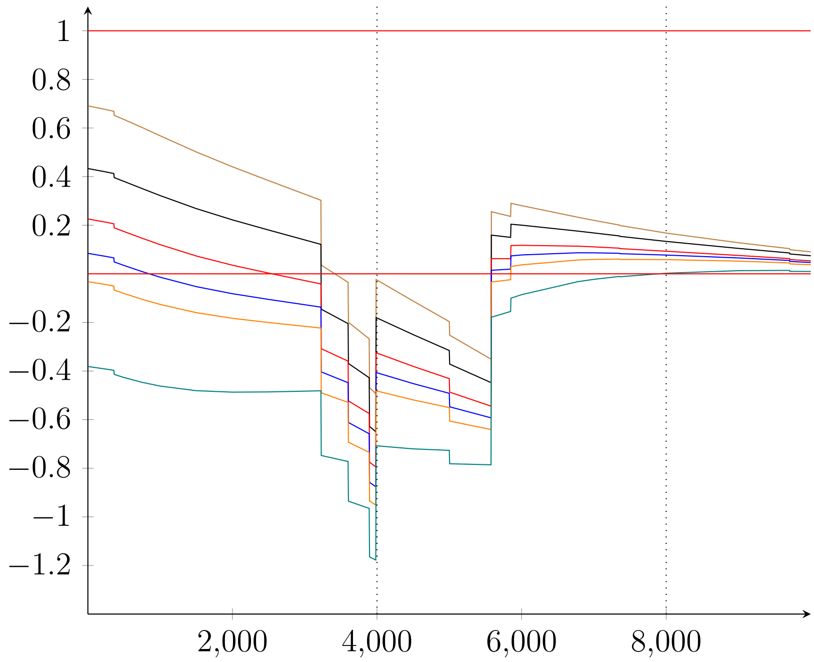


Figure 8: Pareto (in)efficiency of the 1974 US tax and transfer system, varying extensive-margin elasticities.

Notes: Figure 8 shows the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$ before the 1975 reform as a function 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). As in Figure 4, the intensive-margin elasticity is fixed at 0.33. Extensive-margin elasticities are assumed to be constant at 0 (brown line) or 0.4 (red line), to fall from 0.3 at low incomes to 0.1 at higher incomes (black line), to fall from 0.6 to 0.35 (blue line), to fall from 0.7 to 0.45 (orange line), and to fall from 1 to 0.75 (teal line).

Source: Authors' calculations (see Appendix B.1 for details).

5 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, both in the theory of taxation and for the practical design of tax reforms. Our theoretical results show that such reforms can make every one better off, even if no simple one-bracket tax reform can. Our study of the EITC shows that such reforms have also been successfully used in practice. Moreover, we provide sufficient statistics formulas that make it possible to identify such reforms.

That said, there is no reason to believe that two-bracket reforms are only used when this is in everyone's interest. When there are winners and losers from such reforms, the political economy implications (Is the reform politically feasible? Do the winners constitute a majority?) and the implications for social welfare (Do the winners' gains outweigh the losers' losses?) are central issues. In the Online-Appendix of this paper, we develop the tools to analyze them. These tools were not needed, however, for our analysis of the introduction of the EITC in the US. This reform went into a Pareto-improving direction, i.e., there were only winners and no losers. We leave a detailed analysis of two-bracket reforms with more interesting political economy and welfare implications to future research.

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.

——— **and M.R. Jones**, "Do EITC Expansions Pay for Themselves? Effects on Tax Revenue and Public Assistance Spending," *Working Paper Series*, 2019.

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 non-linear tax systems," *CRC TRR 190 Discussion Paper No. 236.*, 2020.

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," *Handbook of income Distribution*, 2000, *1*, 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.

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), pp. 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 Hilary 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*," *The Quarterly Journal of Economics*, 05 1996, *111* (2), 605–637.

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

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

Golosov, M., 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?," Working Paper 2019.

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

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, 1 2003, chapter 3, 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.

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 Taxation with Heterogeneous Skills and Elasticities: Structural and Sufficient Statistics Approaches," *THEMA Working Paper 2016/04*, 2016.

——, ——, **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*, November 2014, *28* (4), 77–98.

——, "Sufficient Statistics Revisited," Working Paper, Princeton University, 2018.

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

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, 9 2010, chapter 3, 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.

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," *Handbook of Public Economics*, 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.

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. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, North-Holland, Amsterdam, 1987.

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.

Appendix

A Proofs

A.1 Proof of Theorem 1

Reforms with one bracket. Adapting inequality (6) to the case of a one-bracket reform we find that a reform that involves higher marginal tax rates, so that $\tau_k > 0$, is Pareto-improving if

$$R^s_\tau(\tau_k, \ell_k, y_k) - \ell_k > 0 \,, \tag{A.1}$$

i.e., if marginal revenue gains are so large that even those whose tax bill increases by the maximal amount of $\ell_k = \max_y \; h_k(y)$ are made better off. Analogously, a reform that involves lower marginal tax rates is Pareto-improving if

$$R^s_\tau(\tau_k, \ell_k, y_k) < 0 \,, \tag{A.2}$$

so that a tax cut leads to larger tax revenues, a logic familiar from analyses of the Laffer curve. We are particularly interested in the question whether a small departure from the status quo can be Pareto-improving. This requires that, for some $\ell_k \geq 0$,

$$R^s_\tau(0, \ell_k, y_k) - \ell_k > 0 \,, \tag{A.3}$$

or

$$R^s_\tau(0, \ell_k, y_k) < 0 \,. \tag{A.4}$$

Lemma A.1

(i) If $R^s_{\tau\ell}(0, 0, y_k) - 1 > 0$, there exists $\ell_k > 0$ so that (A.3) holds.

(ii) If $R^s_{\tau\ell}(0, 0, y_k) < 0$, there exists $\ell_k > 0$ so that (A.4) holds.

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

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

Reforms with two brackets. Lemma A.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.

Lemma A.2 *Consider a two-bracket reform (τ_j, h_j) where h_j is given by (8). Denote by $R^{j2}(\tau_j, \ell_j)$ the change in tax revenue due to a joint reform with two brackets.*

(i) There is a reform (τ_j, h_j) with $\tau_j > 0$ that is Pareto-improving if

$$R^{j2}_{\tau\ell}(0,0) - \max\ \{0, \tau_1\, \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\} \geq 0\ .$$

(ii) There is a reform (τ_j, h_j) with $\tau_j < 0$ that is Pareto-improving if

$$R^{j2}_{\tau\ell}(0,0) - \min\ \{0, \tau_1\, \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\} \leq 0\ .$$

Proof of Lemma A.2. We first note that $R^{j2}_{\tau}(0,0) = 0$.[32] We also note that the function h_j can be written as

$$h_j(y) = \begin{cases} 0, & \text{if} \quad y \leq y_1\ , \\ \tau_1\ (y_1 - y), & \text{if} \quad y \in (y_1, y_1 + \ell_j\, \ell_1]\ , \\ \tau_1\, \ell_j\, \ell_1, & \text{if} \quad y \in (y_1 + \ell_j\, \ell_1, y_2]\ , \\ \tau_1\, \ell_j\, \ell_1 + \tau_2\ (y - y_2), & \text{if} \quad y \in (y_2, y_2 + \ell_j\, \ell_2)\ , \\ \ell_j(\tau_1\, \ell_1 + \tau_2\, \ell_2), & \text{if} \quad y \geq y_2 + \ell_j\, \ell_2\ . \end{cases} \tag{A.5}$$

Hence,

$$\max_y\ h_j(y) \quad = \quad \ell_j\ \max\ \{0, \tau_1 \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\}\ . \tag{A.6}$$

and

$$\min_y\ h_j(y) \quad = \quad \ell_j\ \min\ \{0, \tau_1 \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\}\ . \tag{A.7}$$

Therefore, at $\ell_j = 0$, $R^{j2}_{\tau}(0,0) - \max_y\ h_j(y) = 0$ and $R^{j2}_{\tau}(0,0) - \min_y\ h_j(y) = 0$. Hence, if at $\ell_j = 0$,

$$\begin{aligned} &\tfrac{d}{d\ell_j} \left\{ R^{j2}_{\tau}(0, \ell_j) -\ \ell_j\ \max\ \{0, \tau_1\, \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\} \right\} \\ &= R^{j2}_{\tau\ell}(0,0) - \max\ \{0, \tau_1\, \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\} \\ &> 0\ , \end{aligned}$$

then there exists $\delta > 0$ so that $R^{j2}_{\tau}(0, \ell_j) - \max\limits_y\ h_j(y)$ turns positive as we move from $\ell_j = 0$ to $\ell_j = \delta$. Analogously, if

$$R^{j2}_{\tau\ell}(0,0) - \min\ \{0, \tau_1\, \ell_1, \tau_1\, \ell_1 + \tau_2\, \ell_2\} < 0\ ,$$

then there exists $\delta > 0$ so that $R^{j2}_{\tau}(0, \ell_j) - \min\limits_y\ h_j(y)$ turns negative as we move from $\ell_j = 0$ to $\ell_j = \delta$.

[32] The Gateaux differential is linear in h. Consequently, for any ℓ_j,

$$R^{j2}_{\tau}(0, \ell_j) = \tau_1\ R^{s}_{\tau}(0, \ell_j\, \ell_1, y_1) + \tau_2\ R^{s}_{\tau}(0, \ell_j\, \ell_2, y_2)\ ,$$

Hence,

$$R^{j2}_{\tau}(0,0) \quad = \quad \tau_1\ R^{s}_{\tau}(0,0,y_1) + \tau_2\ R^{s}_{\tau}(0,0,y_2) \quad = \quad 0\ ,$$

where the second inequality follows from the assumption that $R^{s}_{\tau}(0,0,y_k) = 0$, for all y_k.

By the envelope theorem, see equation (3), if $R_{\tau}^{j2}(0, \ell_j) - \max_y \ h_j(y)$, then a reform that involves a small increase of τ_j is Pareto-improving. Analogously, if $R_{\tau}^{j2}(0, \ell_j) - \min_y \ h_j(y) < 0$, then a small decrease of τ_j yields a Pareto improvement.

□

We now turn to the question how these conditions for the existence of a Pareto-improving two-bracket reform relate to the conditions for single-bracket reforms in Lemma A.1. Is it possible that the combination of two reforms – each of which would not be Pareto-improving on a stand alone basis – yields a Pareto improvement. Lemma A.3 provides an answer to this question.

Lemma A.3 *There is a two-bracket reform – i.e., a reform (τ_j, h_j) with h_j given by (8) – that is Pareto-improving if there are income levels y_1 and $y_2 > y_1$ so that*

$$R_{\tau\ell}^{s}(0, 0, y_1) < R_{\tau\ell}^{s}(0, 0, y_2) \, .$$

Proof of Lemma A.3. Fix two income levels y_1 and y_2 with $y_2 > y_1$ and let $R_{\tau\ell}(0, 0, y_1) < R_{\tau\ell}(0, 0, y_2)$. We now construct a Pareto-improving two-bracket reform. To this end, let $\tau_1 < 0$ and $\tau_2 > 0$. Also let $\tau_1 \, \ell_1 + \tau_2 \, \ell_2 = 0 > \tau_1 \, \ell_1$. Then, using equation (A.8),

$$R_{\tau\ell}^{j2}(0, 0) - \max \ \{0, \tau_1 \, \ell_1, \tau_1 \, \ell_1 + \tau_2 \, \ell_2\} = \tau_2 \, \ell_2 \Big(R_{\tau\ell}^{s}(0, 0, y_2) - R_{\tau\ell}^{s}(0, 0, y_1) \Big) > 0 \, .$$

By Lemma A.2, this implies that there exists a Pareto-improving two-bracket reform.

Suppose that $R_{\tau\ell}^{s}(0, 0, y_1), R_{\tau\ell}^{s}(0, 0, y_2) \in [0, 1]$. Then there is no Pareto-improving single-bracket reform for incomes close to y_1 or close to y_2. If $R_{\tau\ell}^{s}(0, 0, y_1) < R_{\tau\ell}^{s}(0, 0, y_2)$ there is still scope for a Pareto improvement, but for one that involves two brackets.

A.2 Proof of Proposition 1

We first show that statement 1 implies statement 2. Suppose that the function $y \mapsto R_{\tau\ell}^{s}(0, 0, y)$ is increasing over some range. Hence, there exist y_1 and $y_2 > y_1$, so that $R_{\tau\ell}^{s}(0, 0, y_1) < R_{\tau\ell}^{s}(0, 0, y_2)$.

Consider the following construction of a two-bracket reform: lower marginal tax rates in a first bracket starting at y_1, i.e., let $\tau_1 < 0$ and increase marginal tax rates in a second bracket starting at y_2, $\tau_2 > 0$. Moreover, choose ℓ_1 and ℓ_2 so that $\tau_1 \ell_1 + \tau_2 \ell_2 = 0$ (while $\ell_1 < y_2 - y_1$). This implies that

$$\max_y \quad h_j(y) \quad = \quad 0 \, ,$$

i.e., no one's tax burden goes up: people with incomes smaller than y_1 have the same tax burden as before as $h_j(y) = 0$ for $y \leq y_1$, people with income between y_1 and y_2 have a lower tax burden since $h_j(y) < 0$ for $y \in (y_1, y_2)$, and people with income above y_2 also have the same tax burden as before, $h_j(y) = 0$ for $y \geq y_2$. Thus, the reform is Pareto-improving if overall revenue goes up. To see that this is the case, note that[33]

$$R_{\tau\ell}^{j2}(0, 0) = \tau_1 \ell_1 \ R_{\tau\ell}^{s}(0, 0, y_1) + \tau_2 \ell_2 \ R_{\tau\ell}^{s}(0, 0, y_2) \, . \tag{A.8}$$

[33] Again the Gateaux differential is linear in h. Consequently, for any ℓ_j,

$$R_{\tau}^{j2}(0, \ell_j) = \tau_1 \ R_{\tau}^{s}(0, \ell_j \ \ell_1, y_1) + \tau_2 \ R_{\tau}^{s}(0, \ell_j \ \ell_2, y_2) \, ,$$

With $\tau_1 \ell_1 + \tau_2 \ell_2 = 0$, we have

$$R^{j2}_{\tau\ell}(0,0) = \tau_2\ \ell_2 \Big(R^s_{\tau\ell}(0,0,y_2) - R^s_{\tau\ell}(0,0,y_1) \Big) > 0\ .$$

Thus, if the revenue effect is larger at the larger level of income, one can phase-in lower marginal tax rates at low income, have an offsetting phase-out at a larger level of income and thereby generate positive revenue – even though no one gets a higher tax bill.

It remains to be shown that statement 2 implies statement 1. By the preceding argument, if a Pareto improvement can be realized by a symmetric two-bracket reform with a tax credit, then it must be the case that

$$R^{j2}_{\tau\ell}(0,0) = \tau_2\ \ell_2 \Big(R^s_{\tau\ell}(0,0,y_2) - R^s_{\tau\ell}(0,0,y_1) \Big) > 0\ .$$

With $\tau_2 \ell_2 > 0$, this implies that $R^s_{\tau\ell}(0,0,y_2) > R^s_{\tau\ell}(0,0,y_1)$. Hence, the function $y \mapsto R^s_{\tau\ell}(0,0,y)$ is increasing over some range.

A.3 Proof of Theorem 2

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

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

of one-bracket reforms. As before, the parameters (τ_j, ℓ_j) determine the size of the reform and the overall revenue is denoted by $R^{jm}(\tau_j, \ell_j)$. The following Lemma states sufficient conditions for the existence of a Pareto-improving reform with m brackets. It is a generalization of Lemma A.2. We omit a formal proof.

Lemma A.4 *Consider a collection $\{(y_k, \tau_j\ \tau_k, \ell_j\ \ell_k)\}_{k=1}^m$ of simple reforms. Let $\tau_0\ l_0 = 0$.*

i) There is a reform (τ_j, h_j) with $\tau_j > 0$ that is Pareto-improving if

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

ii) There is a reform (τ_j, h_j) with $\tau_j < 0$ that is Pareto-improving if

$$R^{jm}_{\tau\ell}(0,0) - \min_{j \in \{0,1,\ldots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k \quad < \quad 0\ .$$

The Lemma 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 Proposition shows that the conditions of the Lemma are indeed violated, with the implication that there is no small Pareto-improving m-bracket reform if the conditions in Theorem 2 hold.

Hence, for any ℓ_j,

$$R^{j2}_{\tau\ell}(0,\ell_j) = \tau_1\ \ell_1\ R^s_{\tau\ell}(0,\ell_j\ \ell_1, y_1) + \tau_2\ \ell_2\ R^s_{\tau\ell}(0,\ell_j\ \ell_2, y_2)\ ,$$

and (A.8) follows from this equation by setting $\ell_j = 0$.

Lemma A.5 *Suppose that the function $y \mapsto R^s_{\tau\ell}(0,0,y)$ is non-increasing, bounded from below by* 0 *and bounded from above by 1. Consider a collection $\{(y_k, \tau_j\ \tau_k, \ell_j\ \ell_k)\}_{k=1}^m$. Let $\tau_0\ l_0 = 0$. Then,*

$$R^{jm}_{\tau\ell}(0,0) - \max_{j\in\{0,1,\dots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k \quad \leq \quad 0\ , \tag{A.9}$$

and

$$R^{jm}_{\tau\ell}(0,0) - \min_{j\in\{0,1,\dots,m\}} \quad \sum_{k=0}^{j} \tau_k\ \ell_k \quad \geq \quad 0\ . \tag{A.10}$$

Proof of Lemma A.5. We first note that the linearity of the Gateaux differential in the direction of reform implies

$$R^{jm}_{\tau\ell}(0,0) \quad = \quad \sum_{k=1}^{m} \tau_k\ \ell_k\ R^s_{\tau\ell}(0,0,y_k)\ . \tag{A.11}$$

We show that the inequality in (A.9) holds for any collection $\{(y_k, \tau_j\ \tau_k, \ell_j\ \ell_k)\}_{k=1}^m$, with $m \geq 2$. The proof that (A.10) holds proceeds along similar lines and is omitted.

Let $j^* = \text{argmax}_l \quad \sum_{k=0}^{l} \tau_k\ \ell_k$. Note that, if j^* is a maximizer, then $\sum_{k=z}^{j^*} \tau_k\ \ell_k \geq 0$ for any $z \in \{0,\dots,l^*\}$ and $\sum_{k=j^*+1}^{z} \tau_k\ \ell_k \leq 0$ for any $z \in \{l^*+1,\dots,m\}$.

As a preliminary step, we verify the following claim: Suppose that $j^* > 0$ and that

$$\sum_{k=z}^{j^*} \tau_k\ \ell_k\ R^s_{\tau\ell}(0,0,y_k) \quad \leq \quad R^s_{\tau\ell}(0,0,y_z) \sum_{k=z}^{j^*} \tau_k\ \ell_k \tag{A.12}$$

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

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

Condition (A.12) is obviously satisfied for $z = j^*$. Hence, a repeated application of the preceding argument yields

$$\sum_{k=1}^{j^*} \tau_k\ \ell_k\ R^s_{\tau\ell}(0,0,y_k) \quad \leq \quad \underbrace{R^s_{\tau\ell}(0,0,y_1)}_{\in(0,1)} \underbrace{\sum_{k=1}^{j^*} \tau_k\ \ell_k}_{\geq 0} \quad \leq \quad \sum_{k=1}^{j^*} \tau_k\ \ell_k \quad = \quad \sum_{k=0}^{j^*} \tau_k\ \ell_k\ . \tag{A.13}$$

An analogous argument implies that

$$\sum_{k=j^*+1}^{m} \tau_k\ \ell_k\ R^s_{\tau\ell}(0,0,y_k) \quad \leq \quad R^s_{\tau\ell}(0,0,y_m) \underbrace{\sum_{k=j^*+1}^{m} \tau_k\ \ell_k}_{\leq 0} \quad \leq \quad 0\ . \tag{A.14}$$

Together (A.13) and (A.14) imply that

$$R^{jm}_{\tau\ell}(0,0) = \sum_{k=1}^{j^*} \tau_k \, \ell_k \, R^s_{\tau\ell}(0,0,y_k) + \sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, R^s_{\tau\ell}(0,0,y_k) \leq \sum_{k=0}^{l^*} \tau_k \, \ell_k \; , \tag{A.15}$$

which proves (A.9). Note that the cases $j^* = 0$ and $j^* = m$ are also covered: with $j^* = 0$, $\sum_{k=1}^{j^*} \tau_k \, \ell_k \, R^s_{\tau\ell}(0,0,y_k)$ does not enter the chain of inequalities and, by the arguments above, $\sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, R^s_{\tau\ell}(0,0,y_k) = 0$. With $j^* = m$, $\sum_{k=j^*+1}^{m} \tau_k \, \ell_k \, R^s_{\tau\ell}(0,0,y_k)$ does not enter.

B Empirical analysis

B.1 Data description and sensitivity analysis

Table B.1 below depicts the sources we use for computing the US tax and transfer systems and for estimating the income distributions of single parents and childless singles in the years 1974 to 1978. We provide further details on our calculations subsequently.

Status quo tax function: the US tax and transfer system in 1974. 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 called Food Stamps). ADFC was available only for single parents and varied to some extent across US states. SNAP was a federal program that was available both for single parents and childless singles, but more generous for single parents. Programs for single parents also depended on the number of children. In the following, we focus on the largest subgroup of single parents: those with two children. In our data, both the modus and the median number of children in single-parent households was two, and the arithmetic mean was close to two. For the years 1975 and later, we also account for the Earned Income Tax Credit.

Income distribution. We estimate the 1974 and 1975 income distributions based on data from the March 1975 and March 1976 Current Population Survey (CPS), respectively. 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. For the benchmark analysis reported in the main text, we estimate the earned income distribution based on the set of single parents with any number of children (see robustness check for single parent with two children below).

In line with the EITC rules, we consider as earned income the sum of (self-reported) wage income and self-employment income. In this sample, 30.9% of single parents and 14.8% of childless singles have zero or negative incomes, while 47% of single parents and 42% of childless singles had strictly positive incomes below 8,000 USD, i.e., in the range of the EITC. 99.1% of these households were also eligible for the EITC if we take into account the capital income test (see further details below). For our benchmark analysis, we estimate the distributions of earned income for both groups using a non-parametric kernel density estimation. Figure B.1 shows the pdf and the cdf of this distribution for single parents and childless singles.

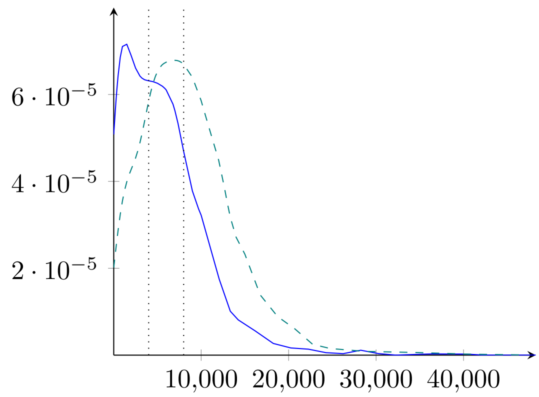
Table B.1: Sources for US tax-transfer system and income data, 1974-1978

Information	Years	Sources
Income tax	1974-1976	Internal Revenue Service, "Instructions for Form 1040", years 1974, 1975, 1976. Accessible for 1974 at https://www.irs.gov/pub/irs_prior/i1040--1974.pdf; for 1975 at www.irs.gov/pub/irs_prior/i1040--1975.pdf; for 1976 at www.irs.gov/pub/irs_prior/i1040--1976.pdf.
	1977-1978	Internal Revenue Service, "Individual income tax returns", years 1977, 1978. Accessible for 1977 at https://www.irs.gov/pub/irs-soi/77inar.pdf; for 1978 at www.irs.gov/pub/irs-soi/78inar.pdf.
EITC	1975-1978	Tax Policy Center, "Earned Income Tax Credit Parameters, 1975-2018", 2018. Accessible at https://www.taxpolicycenter.org/file/178859.
AFDC	1974-1978	Office of the Assistant Secretary for Planning and Evaluation, "Aid to Families with Dependent Children. The Baseline", 1998. URL: https://aspe.hhs.gov/basic-report/aid-families-dependent-children-baseline
	1974	US Department of Health, Education and Welfare, "Aid to Families with Dependent Children: Standards for Basic Needs", 1974. URL: https://hdl.handle.net/2027/mdp.39015088906634.
	1975-1978	TRIM3 project, "TRIM3 AFCD Rules". Accessible at trim3.urban.org.
SNAP	1974-1975	US Bureau of the Census, "Characteristics of Households Purchasing Food Stamps. Current Population Reports", Series 9-23, No. 61, 1976. URL: https://www2.census.gov/library/publications/1976/demographics/p23-061.pdf.
	1976-1978	Federal Register Vol. 41, No. 90, May 7, 1976; Vol. 42, No. 85, May 3, 1977; Vol. 43, No. 95, May 16, 1978. Accessible at https://www.govinfo.gov/app/collection/fr.
Income data	1974-1978	Current Population Survey, Annual Social and Economic Supplement (March waves) 1975 to 1979. Accessible at https://cps.ipums.org/cps/.

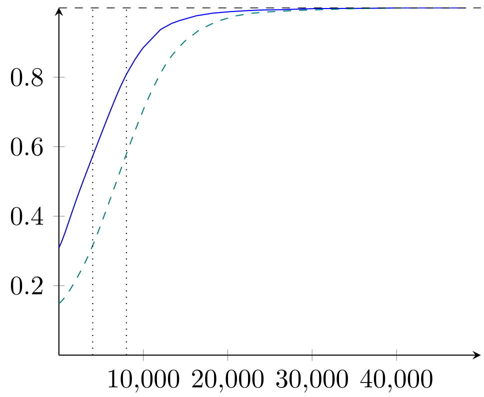
As a robustness check, we also fitted gamma and lognormal distributions to the earned income data in the March 1975 CPS. Figure B.2 shows versions of our test function that results for our estimates based on these parametric distributions. This choice has no substantial effects on our results.

For the benchmark analysis reported in the main text, we estimate the earned income distribution based on the set of single parents with any number of children ($N = 1,494$). We also considered non-parametric kernel estimates of the income distribution for the smaller sample of single parents with exactly two children ($N = 453$). Figure B.3 shows that our results are almost identical in both cases.

Behavioral responses to taxation. We draw on a rich literature providing estimates of labor supply responses at the intensive and the extensive margin – see the discussions in



(a) Probability density functions



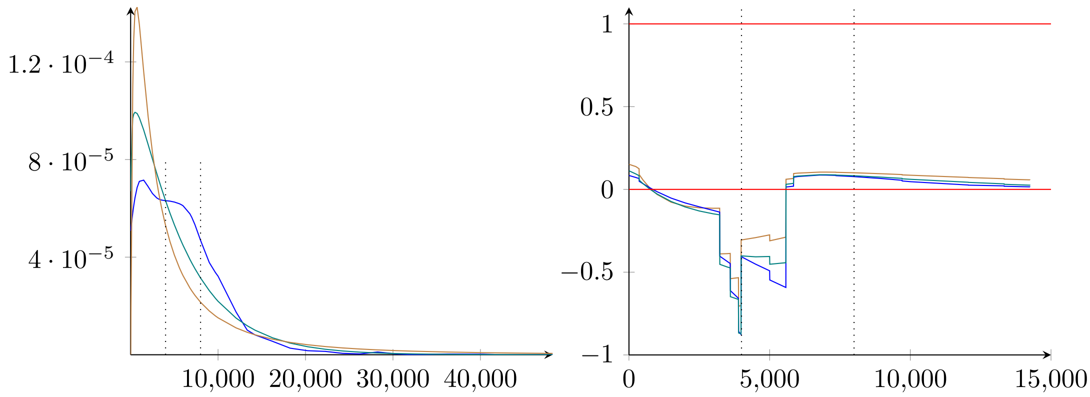
(b) Cumulative distribution functions

Figure B.1: Income distributions of single parents and childless singles in the US 1974.

Notes: Figure B.1 shows the kernel estimates of the US income distributions among single parents (solid blue lines) and childless singles (dashed teal lines) in 1974. Panel (a) depicts the estimated probability density functions; panel (b) depicts the cumulative distribution functions of the income distributions. 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 Table B.1 in Appendix B.1 for details).

Saez, Slemrod and Giertz (2012) or Chetty et al. (2013). Robust findings in the literature are that single mothers and married females respond more strongly at the extensive margin than other subgroups, while married males respond less strongly at both margins. Individuals with little formal education and low incomes respond more strongly at the extensive margin – see, e.g., Juhn, Murphy and Topel (1991), Juhn, Murphy and Topel (2002), Meghir and Phillips (2010). Based on a meta-study and focusing on population-wide averages, Chetty et al. (2013) suggest an intensive-margin elasticity of labor supply with respect to the net-of-tax rate of 0.33, and an extensive-margin elasticity with respect to net labor income of 0.25. Bargain, Dolls, Neumann, Peichl and Siegloch (2014) provide similar estimates for a sample of childless singles. For single parents, various studies suggest an extensive-margin elasticity around or even beyond 0.4, whereas Kleven (2019) obtains estimates close to zero. For the 1975 introduction, Bastian (2020) finds a participation elasticity of 0.58 for his sample of single mothers. Our conclusions on the introduction of the EITC do not depend on this choice, however. Initially, we present results for an intensive-margin elasticity of 0.33, both for single parents and childless singles. Extensive-margin elasticities are assumed to be decreasing with income in both groups, according to the function $\pi_0(y) = \pi_a - \pi_b \ (y/\tilde{y})^{1/2}$, where $\tilde{y}$ equals 50,000 USD. We explore different alternatives for the parameters π_a and π_b. For single parents, we initially assume that the participation elasticity falls from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD (i.e., $\pi_a = 0.6$, $\pi_b = 0.25$), giving rise to an average value of 0.52. For childless singles, we assume π to fall from 0.4 to 0.1, giving rise to an average value around 0.25.

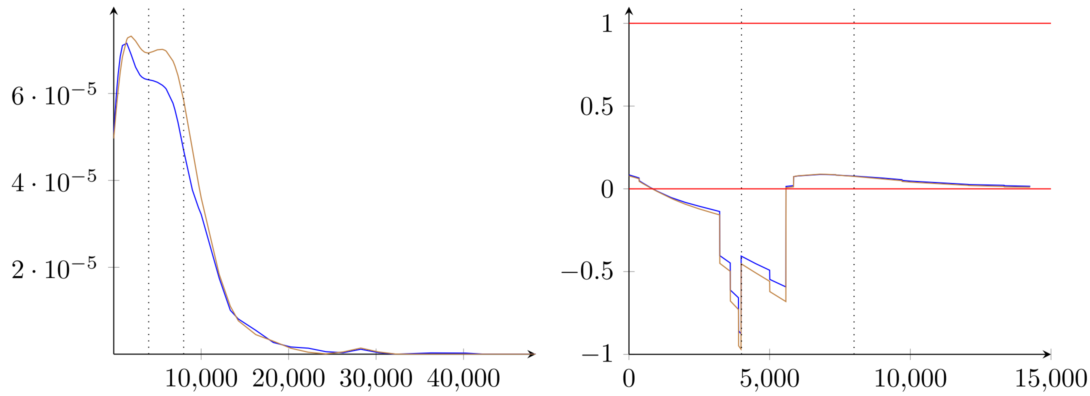


(a) Distributions of earned income (b) Marginal revenue functions

Figure B.2: Pareto (in)efficiency of the 1974 US tax and transfer system for alternative estimates of the income distribution of single parents

Notes: Figure B.2a shows alternative estimates of the income distributions for single parents in 1974: kernel estimation (blue, as in main text), gamma distribution (teal), lognormal distribution (brown). Figure B.2b shows the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$, based on the kernel estimation (blue, as in Figure 4), gamma distribution (teal), lognormal distribution (brown). 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). The intensive-margin elasticity is fixed at 0.33 and the extensive-margin elasticity is assumed to be decreasing with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52.

Source: Authors' calculations (see Table B.1 in Appendix B.1 for details).



(a) Distributions of earned income (b) Marginal revenue function $\mathcal{R}^s_{sp}(y)$

Figure B.3: Pareto (in)efficiency of the 1974 US tax-transfer system for different samples

Notes: Figure B.3a shows the pdfs of the distributions of earned income based on the samples of all single parent (blue) versus single parents with exactly two children (brown line) on the basis of the March 1975 CPS. Figure B.3b depicts the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$, based on the estimated income distributions for all single parents (blue) versus single parents with exactly two children (brown). The former case is considered in the main text, e.g., in Figure 4. 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). The intensive-margin elasticity is fixed at 0.33 and the extensive-margin elasticity is assumed to be decreasing with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52.

Source: Authors' calculations (see Table B.1 in Appendix B.1 for details).

Capital income and assets tests. The following paragraph discusses the relevance of assets and capital income tests, which we have ignored in the main text. We find that they do not affect our main insights. First, single parents were only eligible for the EITC if the sum of their earned income and capital income was below $8,000$ USD. In the March 1975 CPS data, only 0.9% of the single parents with incomes in the EITC range lost eligibility due to high capital incomes (i.e., had an earned income below $8,000$ USD, but a sum of earned income and capital income that exceeded $8,000$ USD). Second, only families with liquid assets below $1,500$ were eligible for welfare transfers (AFDC, SNAP). Following Giannarelli (1992), we approximate liquid assets from the capital income information provided in the CPS data by assuming that assets had a rate of return of 6% for all households. Based on this approximation, we find that about 8% of single parents with earned incomes below the relevant thresholds failed the programs' assets test. Hence, we find that about 7% of single parents were eligible for the EITC, but not for welfare programs.

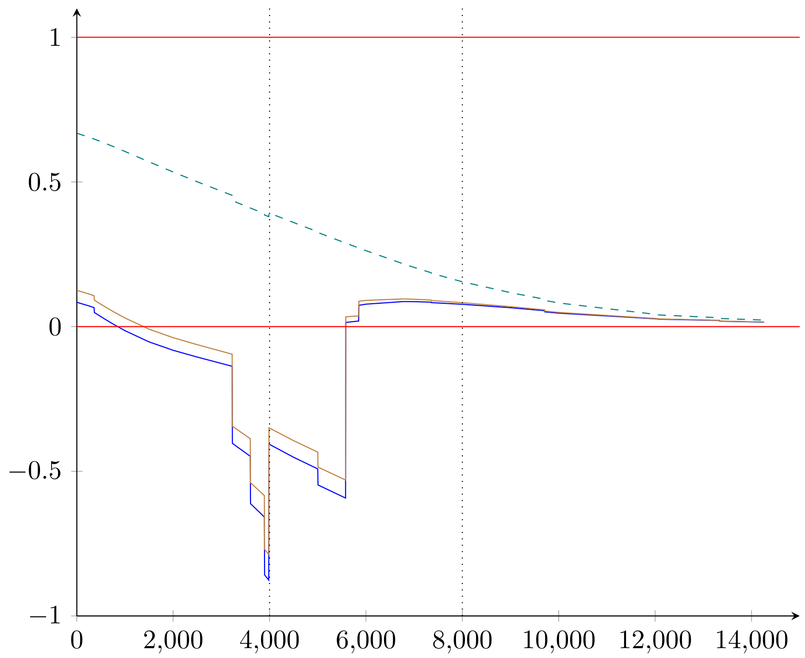


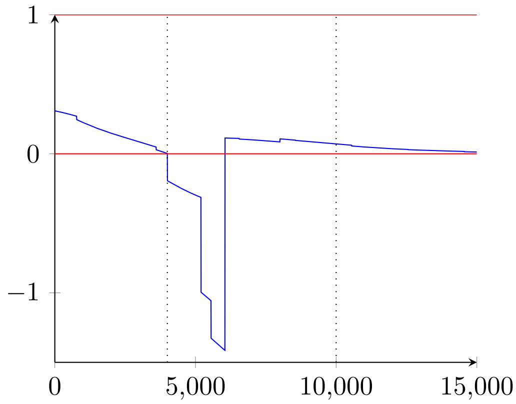
Figure B.4: Pareto (in)efficiency of the 1974 US tax-transfer system, accounting for wealth test

Notes: Figure B.3b depicts the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$ (a) ignoring the wealth test for AFDC and SNAP eligibility (blue line) and (b) taking into account the wealth test (brown line). The former case is considered in the main text, e.g., in Figure 4. The dashed teal line shows $\mathcal{R}^s_{sp}(y)$ ignoring AFDC and SNAP, i.e., based on the income tax schedule alone. 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). The intensive-margin elasticity is fixed at 0.33 and the extensive-margin elasticity is assumed to be decreasing with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, with average value of 0.52.
Source: Authors' calculations (see Table B.1 in Appendix B.1 for details).

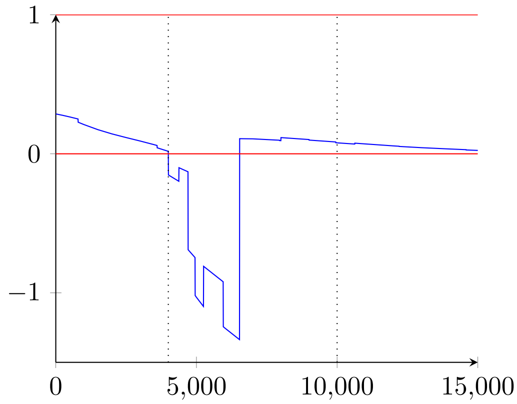
Our benchmark analysis in the main text ignores the assets and capital income tests (see, e.g., Figure 4). This is the appropriate procedure to determine whether the 1974 US tax and transfer system was efficient or not: The US authorities could have introduced an EITC with the same assets test as in place for AFDC and SNAP. Figure 4 demonstrates the existence of Pareto-improving reforms with this feature. Alternatively, we can ask whether the EITC reform with its specific capital income test (differing from the AFDC assets test) was Pareto-improving. To answer this question, the brown line in Figure B.4 below depicts a version of our test function $\mathcal{R}^s_{sp}(y)$ – giving the revenue implications of small one-bracket reforms at any income y – under the assumption that 7% of single parents were not eligible for AFDC and SNAP, while all other single parents were eligible. For comparison, the blue line depicts the (benchmark) marginal revenue function that applies if all single parents are eligible for welfare, and the dashed teal line depicts the marginal revenue function that would apply if none of the single parents was eligible for welfare. As Figure B.4 shows, taking into account the wealth test leaves our results qualitatively unchanged: There was large scope for Pareto-improvements both by means of one-bracket and two-bracket reforms.

B.2 Results for further post-reform years

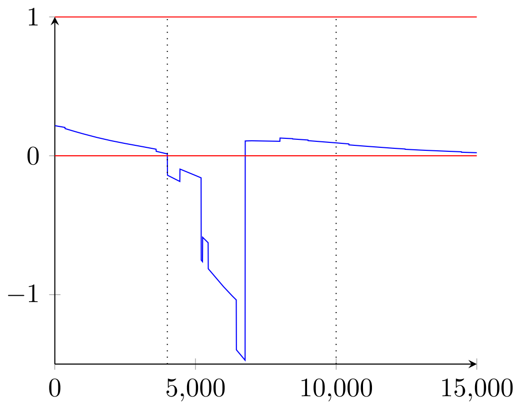
Figure B.5 depicts the marginal revenue function for all years between the EITC introduction in 1975 and its first expansion in 1979. It shows that the inefficiencies in the US tax and transfer system remained during this time, although the introduction of the EITC in 1975 narrowed the income range on which the conditions for Pareto efficiency are violated (relative to the 1974 tax and transfer system). From 1975 to 1978, however, the income range with inefficiencies increased again to cover almost the entire phase-out range between $4,000$ and $8,000$ USD. In 1979, US authorities expanded the EITC: They extended the phase-in range to include all income levels up to $5,000$ USD, introduced a plateau range between $5,000$ and $6,000$ USD and adjusted the phase-out range to run from $6,000$ to $10,000$ USD (see dotted vertical lines in Figure B.5).



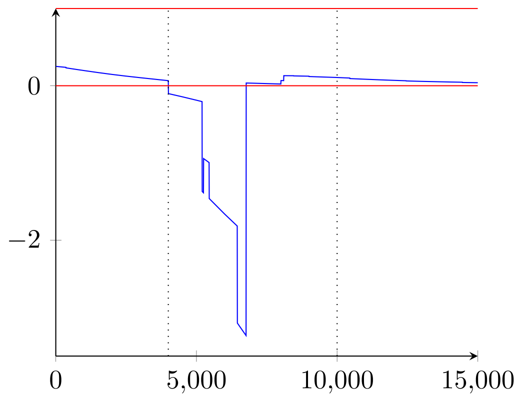
(a) 1975.



(b) 1976.



(c) 1977.



(d) 1978.

Figure B.5: Pareto (in)efficiency of the US tax and transfer system for post-reform years 1975-1978

Notes: Figure B.5 shows the marginal revenue function for single parents $\mathcal{R}^s_{sp}(y)$ for the years 1975 to 1978 after the EITC introduction. In 1979, the EITC was extended in the income range between $4,000$ USD (first dotted line) and $10,000$ USD (second dotted line). The intensive-margin elasticity is fixed at 0.33 and extensive-margin elasticity is assumed to be decreasing with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD, giving and average value of 0.52.

Source: Authors' calculations (see Table B.1 in Appendix B.1 for details).

Online Appendix (not for publication)

C Reforms with winners and losers

If a reform is not Pareto-improving, then there are individuals who are made better off and others who are made worse off. This raises the question whether the reform is worthwhile from a welfare perspective, i.e., whether the gains of the winners exceed the costs of the losers. From a political economy perspective, the question is whether the winners form a majority. In this section, we provide answers to these questions. We characterize the welfare and the political economy implications of reforms with one or two brackets.

Who are the winners? Who are the losers? Whether or not a person benefits from intensifying a reform follows from equation (3): a person benefits if and only if $R_\tau(\tau, h) - h(y^*(\cdot)) > 0$.

For reforms with only one bracket, h is a non-decreasing function. Consider a reform that involves an increase of marginal tax rates by $\tau > 0$ and which is neither Pareto-improving nor Pareto-damaging. Then, there must exist a cutoff level of income $\hat{y}$ such that all individuals with an income below are beneficiaries and all individuals with an income above are made worse off.[34]

Now consider a reform with two brackets and suppose that the overall revenue effect is positive, $R_\tau(\tau, h) \geq 0$. Also, suppose that there is a phase-in region starting at income level y_1 that involves lower marginal tax rates, $\tau_1 < 0$, and a phase-out region starting at y_2 involving higher marginal tax rates, $\tau_2 > 0$. If such a reform makes some people worse off, there exists a cutoff level of income $\hat{y} > y_2$ such that people with an income above are worse off and people with an income below are better off. People below the cutoff benefit from the increase of revenue and possibly also from a lowering of their taxes. For people with an income above the cutoff, the increase of their tax burden overturns the positive revenue effect.

If, by contrast, the reform comes with a revenue loss, $R_\tau(\tau, h) < 0$, then there is a cutoff level $\hat{y}_1 \geq y_1$ so that people with an income below $\hat{y}_1$ are made worse off and people with an income above are beneficiaries of the reform. People with an income below the cutoff suffer from the loss of revenue. Some of them also gain from a reduction of their tax burden, but this gain is smaller than the revenue loss. For people with an income above $\hat{y}_1$ the tax cut dominates the revenue loss. Possibly, there exists a second cutoff $\hat{y}_2 > \hat{y}_1$ so that people with an income above $\hat{y}_2$ are also made worse off. If the phase-out goes sufficiently far, then people with a high enough income do not benefit much from tax cuts, if at all, so that the revenue loss is the dominating force.

Welfare implications. With an additive social welfare function, the welfare implications of a tax reform are given by

$$\mathcal{W}(\tau, h) := \mathbf{E}\left[g(\theta)\ v(R(\tau, h), \tau, h, \theta)\right] ,$$

[34] For a reform with $\tau < 0$, these roles are reversed, i.e., the beneficiaries have incomes above the cutoff and the reform losers have incomes below.

where $g : \theta \mapsto g(\theta)$ is a function specifying the welfare weights for different types of individuals. Upon using (3), the welfare implications of a marginal change of τ, can be written as

$$\mathcal{W}_\tau(\tau, h) = \lambda(\tau, h) \left\{ R_\tau(\tau, h) - \frac{1}{\lambda(\tau, h)} \mathbf{E} \left[g(\theta) \; u_c(\cdot) \; h(y^*(\cdot)) \right] \right\} , \tag{C.1}$$

where $\lambda(\tau, h) = \mathbf{E} \left[g(\theta) \; u_c(\cdot) \right]$ is a welfare-weighted population average of the marginal utility of consumption, evaluated at reform intensity τ, for a given reform direction h. The expression $\mathbf{E} \left[g(\theta) \; u_c(\cdot) \; h(y^*(\cdot)) \right]$ by contrast, looks only at people whose tax burden is affected by the reform, i.e., people with $h(y^*(\cdot)) \neq 0$.

Pareto-improving reforms and the inverse tax problem. The canonical optimal tax problem is to find the tax policy that maximizes a given social welfare function. The literature on the *inverse tax problem*, by contrast, takes the tax function as given and tries to find the corresponding welfare function. It exploits first-order conditions from welfare maximization and seeks to identify the welfare weights for which these first-order conditions are satisfied.[35] As we now show, there is a connection between our results on Pareto-improving tax reforms and the inverse tax problem. Specifically, we show that the existence of a Pareto-improving reform implies that the solution to the inverse tax problem give rise to negative welfare weights, and hence contradicts the maximization of a Paretian social welfare function, i.e., of a social welfare function that is non-decreasing in the different types' utilities. Moreover, we show that the pattern in those negative weights is informative about what type of reform would be Pareto-improving.

The welfare implications of a small single-bracket reform at income level y_1 are given by

$$\mathcal{W}^s_{\tau\ell}(0, 0, y_1) = \lambda_0 \Big\{ R^s_{\tau\ell}(0, 0, y_1) - \frac{1}{\lambda_0} \left[1 - F_y(y_1) \right] \; \mathbf{E} \left[g(\theta) \; u_{c0}(\theta) \mid y_0(\theta) \geq y_1 \right] \Big\} ,$$

where $u_{c0}(\theta)$ is a shorthand for the marginal utility of consumption of type θ in the status quo and $\lambda_0 := \mathbf{E} \left[g(\theta) \; u_{c0}(\theta) \right]$ is the corresponding welfare-weighted population average.

The first-order condition of welfare-maximization is that $\mathcal{W}^s_{\tau\ell}(0, 0, y_1) = 0$. Suppose that this first-order condition is satisfied and that there is a possibility of a Pareto-improving tax cut such that $R^s_{\tau\ell}(0, 0, y_1) < 0$. Then,

$$\mathbf{E} \left[g(\theta) \; u_{c0}(\theta) \mid y_0(\theta) \geq y_1 \right] < 0 ,$$

i.e., the average weights on those who would benefit the most from such a tax cut are negative. This is incompatible with the maximization of a Paretian social welfare function. Analogously, if there is a possibility of a Pareto-improving tax raise, $R^s_{\tau\ell}(0, 0, y_1) > 1$, then the complementary set of agents is found to have a negative average weight,

$$\mathbf{E} \left[g(\theta) \; u_{c0}(\theta) \mid y_0(\theta) < y_1 \right] < 0 .$$

We now consider two-bracket reforms. The welfare implications of a small two-bracket

[35] See, e.g., Christiansen and Jansen (1978), Blundell et al. (2009), Bourguignon and Spadaro (2012), Bargain et al. (2011), Jacobs et al. (2017), Lockwood and Weinzierl (2016), or Hendren (2020).

reform can be written as

$$\begin{aligned}
\mathcal{W}^{j2}_{\tau\ell}(0,0) &= \lambda_0 \Big\{ \tau_1\, \ell_1\, R^s_{\tau\ell}(0,0,y_1) + \tau_2\, \ell_2\, R^s_{\tau\ell}(0,0,y_2) \\
&\qquad -\tfrac{1}{\lambda_0}\tau_1\, \ell_1\, [1 - F_y(y_1)] \quad \mathbf{E}\,[g(\theta)\, u_{c0}(\theta) \mid y_0(\theta) \geq y_1] \\
&\qquad -\tfrac{1}{\lambda_0}\tau_2\, \ell_2\, [1 - F_y(y_2)] \quad \mathbf{E}\,[g(\theta)\, u_{c0}(\theta) \mid y_0(\theta) \geq y_2] \Big\} \,.
\end{aligned}$$

If $R^s_{\tau\ell}(0,0,y_2) > R^s_{\tau\ell}(0,0,y_1)$ there is a Pareto-improving two-bracket reform with $\tau_2 > 0$ and $\tau_1\, \ell_1 + \tau_2\, \ell_2 = 0$. If such a reform satisfies the first order condition, $\mathcal{W}^{j2}_{\tau\ell}(0,0) = 0$, then

$$\mathbf{E}\Big[g(\theta)\, u_{c0}(\theta) \mid y_0(\theta) \in [y_1, y_2]\Big] < 0 \,.$$

Again, this implies a contradiction to the maximization of a Paretian social welfare function.

Political support. Consider a reform in direction h. Is there a majority in favor of intensifying it, i.e., in favor of raising τ?

When there is a single cutoff level of income $\hat{y}_1$, dividing those who are made better off and those who are made worse off, then it is easy to provide an answer: Suppose, for concreteness, that those with an income below $\hat{y}_1$ are made better off and those with an income above are made worse off. Then there is majority support for the reform if and only if the median level of income is below the cutoff $\hat{y}_1$. The following Proposition, which we state without proof, contains a more formal version of this observation.

Proposition C.1 *Suppose there is a cutoff level of income $\hat{y}_1$ so that either*

$$R_\tau(\tau, h) - h(y^*(\cdot)) \begin{cases} < 0, & \text{if} \quad y^*(\cdot) < \hat{y}_1 \,, \\ > 0, & \text{if} \quad y^*(\cdot) > \hat{y}_1 \,, \end{cases}$$

or

$$R_\tau(\tau, h) - h(y^*(\cdot)) \begin{cases} > 0, & \text{if} \quad y^*(\cdot) < \hat{y}_1 \,, \\ < 0, & \text{if} \quad y^*(\cdot) > \hat{y}_1 \,. \end{cases}$$

Then the following statements are equivalent:

1. *A majority of individuals benefits if the reform is intensified.*

2. *The individual with median income benefits if the reform is intensified.*

Proposition C.1 generalizes a median voter theorem in Bierbrauer and Boyer (2018) that applies to reforms so that h is a monotonic function. If h is monotonic, then there is a cutoff level of income dividing reform winners and losers. A two-bracket reform with a phase-in range and a phase-out range corresponds to an h function that is not monotonic. Still, as our discussion above has shown, such reforms may also give rise to a single cutoff dividing winners and losers. The median voter result in Proposition C.1 applies also in this case.

We now turn to two-bracket reforms with two cutoffs $\hat{y}_1$ and $\hat{y}_2$ so that individuals with incomes below $\hat{y}_1$ or above $\hat{y}_2$ are made worse off and individuals with an income between $\hat{y}_1$

and $\hat{y}_2$ are made better off. As explained above, reforms with a phase-in range and a phase-out range that come with a loss of overall tax revenue have this property. In particular, this includes symmetric two-bracket tax cuts.

To analyze whether such a reform gets sufficient political support, it is useful to define the midpoint of the plateau between the phase-in and the phase-out range:

$$y_m := \frac{1}{2}(y_1 + \ell_1 + y_2) \ .$$

Individuals with an income sufficiently close to y_m are beneficiaries of the reform and individuals with an income far away are opponents. More formally, there exist d_s and d_o, with the indices standing, respectively, for *s*upport and *o*pposition, so that

$$\mid y^*(\cdot) - y_m \mid \quad < \quad d_s \quad \text{implies} \quad R_\tau(\tau, h) - h(y^*(\cdot)) > 0 \ ,$$

and

$$\mid y^*(\cdot) - y_m \mid \quad > \quad d_o \quad \text{implies} \quad R_\tau(\tau, h) - h(y^*(\cdot)) < 0 \ .$$

Finally, let y_{dM} be the median of the cross-section distribution of $\mid y^*(\cdot) - y_m \mid$. That is, y_{dM} is defined by the property that half of the population has an income that is closer to y_m than y_{DM} and half of the population has an income that is more distant.

Proposition C.2 *Consider a two-bracket reform with a phase-in and a phase-out range. Suppose there are cutoffs $\hat{y}_1$ and $\hat{y}_2$ so that individuals with incomes below $\hat{y}_1$ or above $\hat{y}_2$ are worse off and individuals with an income between $\hat{y}_1$ and $\hat{y}_2$ are better off.*

1. *Suppose that $\mid y_{dM} - y_m \mid < d_s$, then there is majority support for the reform.*

2. *Suppose that $\mid y_{dM} - y_m \mid > d_o$, then there is no majority support for the reform.*

3. *If $d_s = d_o$, there is majority support if and only if there is support by the person with income y_{dM}.*

We omit a formal proof, but explain the logic behind it. There is a neighborhood of y_m so that all individuals with an income in this neighborhood are beneficiaries of the reform. The largest such neighborhood is characterized by the distance d_s. If the person with income y_{dM} is in, then more than fifty percent of the population are in and hence there is majority support for the reform. This is the first statement in the Proposition. The second statement asserts that there is also another distance $d_o \geq d_s$ so that everyone with an income that differs from y_m by more than d_0 is an opponent of the reform. If the person with income y_{dM} is such an opponent, then any one with an income even further away from y_m is also an opponent. Hence, there is a majority of opponents. The distribution of supporters and opponents may not be symmetric around y_m. Hence, there is the possibility that $d_s < d_0$, giving rise to the possibility that, say, some people with an income below y_m are opponents and some people with an equal distance, but an income above y_m are supporters. For the special case of a symmetric two-bracket tax cut with $\ell_1 = \ell_2$ and $\tau_1 = -\tau_2$, we have $d_0 = d_s$ and hence an equivalence of majority support and support by the person with income y_{dM}. This is the third statement of the Proposition.

D Sufficient statistics with behavioral responses at the intensive and the extensive margin

We consider a setting with two-dimensional heterogeneity. Individuals differ both in fixed and variable costs associated with the generation of income. Such a framework has been suggested by the literature that analyzes earnings subsidies from an optimal-tax perspective (see, e.g., Saez (2002), Jacquet et al. (2013), or Hansen (2019)). The analysis in this part of the Appendix yields a characterization of $y \mapsto R^s_{\tau\ell}(0,0,y)$ that depends on labor supply elasticities at the intensive and the extensive margin.

Henceforth, fixed costs are captured by a parameter γ, variable costs by a parameter ω. Thus, we write $\theta = (\omega, \gamma)$ for an individual's type. Variable cost types and fixed cost types belong, respectively, to subsets of the positive reals that we denote by $\Omega = [\underline{\omega}, \overline{\omega}]$ and $\Gamma = [\underline{\gamma}, \overline{\gamma}]$. The joint distribution is denoted by F. The utility that an individual with type (ω, γ) derives from a (c, y)-pair that involves positive earnings is denoted by $u(c, y, \omega, \gamma)$. We denote by $u^{no}(c_0)$ the utility of individuals with no earnings. The function u^{no} is assumed to be increasing and weakly concave.

Variable costs. To capture variable costs, we assume that preferences satisfy the Spence-Mirrlees single-crossing property: Consider two individuals with the same fixed cost type γ, and an arbitrary point in the (c, y)-space with $y > 0$. At any such point, an individual with a higher ω-type has a flatter indifference curve. The interpretation is that, due to her lower variable costs, she needs less compensation for a marginal increase of her earnings. Formally, for any given γ, $\omega' > \omega$ implies that

$$-\frac{u_y(c, y, \omega', \gamma)}{u_c(c, y, \omega', \gamma)} < -\frac{u_y(c, y, \omega, \gamma)}{u_c(c, y, \omega, \gamma)} \; .$$

for any pair (c, y) with $y > 0$.

Let $C : y \mapsto C(y)$ be a non-decreasing function, interpreted as the boundary of a budget set that individuals face. An implication of the Spence-Mirrlees single crossing property is as follows: Consider two individuals who differ only in the variable cost type. If type ω weakly prefers an earnings level y' over an earnings level $y < y'$, then any type $\omega' > \omega$ strictly prefers y' over y. More formally, for any γ, any pair ω', ω with $\omega' > \omega$, and any pair y', y with $y' > y$,

$$u(C(y'), y', \omega, \gamma) \geq u(C(y), y, \omega, \gamma) \quad \text{implies} \quad u(C(y'), y', \omega', \gamma) > u(C(y), y, \omega', \gamma) \; .$$

Fixed costs. Fixed costs affect the compensation that individuals demand for positive earnings. Let $\pi(c, y, \omega, \gamma)$ be such that

$$u(c + \pi(c, y, \omega, \gamma), y, \omega, \gamma) = u^{no}(c_0) \; .$$

We assume that π is an increasing function of γ.[36]

[36]This property holds for various preference specifications that have been explored in the literature. In particular, it holds for separable utility functions of the form $u(c, y, \omega, \gamma) = \tilde{u}(c, y, \omega) - \gamma \, \mathbf{1}_{y>0}$, where $\mathbf{1}$ is the indicator function. It also holds for specifications with monetary fixed costs $u(c, y, \omega, \gamma) = \tilde{u}(c - \gamma \, \mathbf{1}_{y>0}, y, \omega)$. The two classes coincide if the function $\tilde{u}$ is quasi-linear in c.

An implication is as follows: Consider two individuals who differ only in their fixed-cost type γ. Given a non-decreasing consumption schedule $C : y \mapsto C(y)$, if type γ prefers an earnings level of 0 over an earnings level of $y > 0$, then any individual with a fixed-cost type $\gamma' > \gamma$, will also prefer 0 over y; for any ω, any pair γ', γ with $\gamma' > \gamma$, and any $y > 0$,

$$u^{no}(c_0) \geq u(C(y), y, \omega, \gamma) \quad \text{implies} \quad u^{no}(c_0) > u(C(y), y, \omega, \gamma') \ .$$

The Spence-Mirrlees single-crossing property preserves monotonicity of choices in variable costs. For a given continuous consumption schedule $C : y \mapsto C(y)$, let $y^*(\omega, \gamma)$ be the utility-maximizing choice of type (ω, γ). By the Spence-Mirrlees single-crossing property, $\omega' > \omega$ implies $y^*(\omega', \gamma) \geq y^*(\omega, \gamma)$. In particular, $y^*(\omega', \gamma) = 0$ implies $y^*(\omega, \gamma) = 0$. Thus, for any given γ, there is a cutoff type $\hat{\omega}(\gamma)$ so that $\omega < \hat{\omega}(\gamma)$ implies $y^*(\omega, \gamma) = 0$, whereas $\omega \geq \hat{\omega}(\gamma)$ implies $y^*(\omega, \gamma) > 0$.

The earnings function will generally exhibit an upward jump at $\hat{\omega}(\gamma)$. With C continuous, raising y slightly above 0 comes only with a small gain in consumption utility, but an upward jump of effort costs. Thus, a significant increase of earnings is needed to have a gain in consumption utility that offsets these effort costs. Moreover, by our assumption on fixed costs, $\gamma' > \gamma$ implies that $\hat{\omega}(\gamma') \geq \hat{\omega}(\gamma)$.

The earnings function y^* is bounded away from zero for all (ω, γ) with $\omega > \hat{\omega}(\gamma)$. Over this domain, we take y^* to be a non-decreasing function of γ. Thus, $y^*(\omega, \gamma) > 0$, $y^*(\omega, \gamma') > 0$ and $\gamma' > \gamma$ imply that $y^*(\omega, \gamma') \geq y^*(\omega, \gamma)$.[37]

For a given reform direction h, we denote by $y^*(e, \tau, h, \omega, \gamma)$ the solution to the problem

$$\max_{y \geq 0} \ u(c_0 + e + y - T_0(y) - \tau \ h(y), y, \omega, \gamma)$$

and the corresponding indirect utility by $v(e, \tau, h, \omega, \gamma)$. The parameter e stands for a source of income that is exogenous from an individuals' perspective. In the subsequent analysis, e will be equal to the change in tax revenues, $e = R(\tau, h)$. If $y^*(e, \tau, h, \omega, \gamma) = 0$, then $v(e, \tau, h, \omega, \gamma) = u^{no}(c_0 + e)$.

The earnings function y^* exhibits a discontinuity at $\hat{\omega}(\cdot)$. Earnings are zero for types below $\hat{\omega}(\cdot)$ and bounded away from zero for types above. Individuals with type $\hat{\omega}(\cdot)$ are indifferent between earnings of zero and a strictly positive earnings level. It is convenient to assume that these individuals have positive earnings. Thus, we assume that

$$y^*(R(\cdot), \tau, \hat{\omega}(\cdot), \gamma) > 0 \ .$$

Intensive-margin responses. For one-bracket reforms, the derivative of the function y^*_τ with respect to τ gives how earnings respond to small changes in marginal tax rates for incomes that lie in that bracket. These are the behavioral responses at the intensive margin.

[37] For separable utility functions of the form $u(c, y, \omega, \gamma) = \tilde{u}(c, y, \omega) - \gamma \ \mathbf{1}_{y>0}$, $y^*(\omega, \gamma) > 0$, $y^*(\omega, \gamma') > 0$ and $\gamma' > \gamma$ imply that $y^*(\omega, \gamma') = y^*(\omega, \gamma)$. For specifications $u(c, y, \omega, \gamma) = \tilde{u}(c - \gamma \ \mathbf{1}_{y>0}, y, \omega)$ with concave consumption utility, $y^*(\omega, \gamma) > 0$, $y^*(\omega, \gamma') > 0$ and $\gamma' > \gamma$ imply that $y^*(\omega, \gamma') > y^*(\omega, \gamma)$.

Extensive-margin responses. For a given reform direction h, we view the cutoff type $\hat{\omega}$ not only as a function of γ, but also as a function of the size of the reform as measured by τ. Formally, for given γ, the cutoff type $\hat{\omega}(\tau, \gamma)$ is defined as the value of ω that solves

$$u^{no}(c_0 + R(\tau, h)) = v(R(\tau, h), \tau, \omega, \gamma) \ .$$

The effect of a small change of the reform intensity τ on the cutoff type $\hat{\omega}$ is obtained by computing a total differential of this equation. This yields, invoking again the envelope theorem,

$$u_c^{no}(\cdot) \ R_\tau(\tau, h) = u_c(\cdot) \left(R_\tau(\tau, h) - h(\cdot)\right) + u_\omega(\cdot) \ \hat{\omega}_\tau(\tau, \gamma) \ ,$$

where the functions u_c, u_ω and h are evaluated at $y = y^*(R(\cdot), \tau, \hat{\omega}(\cdot), \gamma)$. Equivalently,

$$\hat{\omega}_\tau(\tau, \gamma) = \frac{u_c^{no}(\cdot)}{u_\omega(\cdot)} R_\tau(\tau, h) - \frac{u_c(\cdot)}{u_\omega(\cdot)} \left(R_\tau(\tau, h) - h(\cdot)\right) \ . \tag{D.1}$$

To interpret these expressions, consider the following thought experiment: a fraction $F_\Omega(\hat{\omega}(\tau, \gamma) \mid \gamma)$ of individuals with fixed cost type γ has zero earnings, where $F_\Omega(\cdot \mid \gamma)$ is the distribution of variable cost types ω conditional on the fixed cost type being γ. Consider a small increase of transfers only for the unemployed, the marginal effect on $F_\Omega(\cdot \mid \gamma)$ is given by

$$P^{no}(\tau, \gamma) := f_\Omega(\hat{\omega}(\tau, \gamma) \mid \gamma) \ \frac{u_c^{no}(\cdot)}{u_\omega(\cdot)} \ , \tag{D.2}$$

where the letter P is chosen to indicate a marginal effect on Participation. Alternatively, the effect of a transfer only to those with positive earnings is given by

$$P(\tau, \gamma) = f_\Omega(\hat{\omega}(\tau, \gamma) \mid \gamma) \ \frac{u_c(\cdot)}{u_\omega(\cdot)} \ . \tag{D.3}$$

Thus, assuming that the distribution F and these marginal effects are known is equivalent to assuming that the extensive-margin elasticities and semi-elasticities that are ubiquitous in the related literature are known.

Notation. It is convenient to use a shorthand for endogenous variables at the status quo. For instance, we will occasionally write $\hat{\omega}_0(\gamma) := \hat{\omega}(0, \gamma)$ for the type at the particiaption margin, among those with fixed cost type γ. We write $y_0^*(\omega, \gamma) := y^*(0, 0, h, \omega, \gamma)$ for income in the status quo, and similarly for other variables. Given a fixed-cost type γ, we denote by $\omega_0(y, \gamma)$ the variable cost type who chooses earnings of y in the status quo. If we evaluate partial derivatives at the status quo, we occasionally write $\hat{\omega}_{0\tau}(\gamma)$ or $y_{0e}^*(\omega, \gamma)$ and so on. We, moreover, write $\hat{y}_0^*(\gamma) := y_0^*(\hat{\omega}_0(\gamma), \gamma)$ for the status quo income of the cutoff type among those with fixed costs of γ. Finally,

$$I(\omega_0(y, \gamma) \mid \gamma) = E_\Omega \left[T_0'(y_0^*(s, \gamma)) \ y_{0e}^*(s, \gamma) \mid s \geq \omega_0(y, \gamma), \gamma\right] \ ,$$

is a measure of the size of income effects among those individuals with fixed cost type γ who have earnings exceeding y.

Revenue implications of reforms with one bracket. By our analysis in the previous section, to understand whether a given tax system can be reformed in a Pareto-improving way, we need to check whether the function $y \mapsto R^s_{\tau\ell}(0,0,y)$ is non-increasing, bounded from below by 0 and bounded from above by 1. The following Proposition provides a characterization of this function for the given setup with variable and fixed costs of productive effort.

Proposition D.1 *Suppose that, for any given γ, $\omega_0(y,\gamma)$ is strictly increasing in y, whenever $y > 0$. Also suppose that, for all (ω,γ), income in the status quo satisfies the first-order conditions of utility-maximization whenever $y^*_0(\omega,\gamma) > 0$. Then,*

$$R^s_{\tau l}(0,0,y) = \left(1 - \mathcal{M}^0\right)^{-1} \Big(\mathcal{I}(y) - \mathcal{X}(y)\Big) \, ,$$

where

$$\mathcal{M}^0 = E_\Gamma \Big[(1 - F(\hat{\omega}_0(\gamma) \mid \gamma))\, I(\hat{\omega}_0(\gamma) \mid \gamma) - T_0(\hat{y}^*_0(\gamma))\ \left(P^{no}_0(\gamma) - P_0(\gamma)\right)\Big] \, ,$$

$$\begin{aligned}\mathcal{I}(y) \quad &= \quad T'_0(y)\ E_\Gamma \left[f(\omega_0(y,\gamma) \mid \gamma) \tfrac{y^*_{0\tau}(\omega_0(y,\gamma),\gamma)}{y^*_{0\omega}(\omega_0(y,\gamma),\gamma)}\right] \\ &\qquad + E_\Gamma\left[(1 - I(\omega_0(y,\gamma) \mid \gamma))(1 - F(\omega_0(y,\gamma) \mid \gamma))\right] \, ,\end{aligned}$$

and

$$\mathcal{X}(y) = \int_{\underline{\gamma}}^{\overline{\gamma}} T_0(\hat{y}^*_0(\gamma))\ P_0(\gamma)\ \mathbf{1}(\hat{y}^*_0(\gamma) \geq y)\ f_\Gamma(\gamma) d\gamma \, .$$

The proof of Proposition D.1 can be found below. Proposition D.1 shows that the revenue effect of a small single-bracket reform at income level y, $R^s_{\tau l}(0,0,y)$, can be decomposed into an extensive-margin effect $\mathcal{X}(y)$, an intensive-margin effect $\mathcal{I}(y)$, and a multiplier $\mathcal{M}^0$. The assumptions that $\omega_0(y,\gamma)$ is strictly increasing in y and that individual behavior can be described by first-order conditions are made for ease of exposition. They avoid complications due to bunching.

The extensive-margin effect is shaped by the employment response of those individuals who are close to indifferent between staying out of the labor market and entering. More specifically, $\mathcal{X}(y)$ gives an average for all types who choose earnings of at least y when entering the labor force. A change of the marginal rates in a bracket that begins at y has no effect on individuals who only consider incomes lower than y. For those who consider an income of y or above, there is a negative effect on participation and this tends to lower the revenue that is raised by such a reform. Naturally, the effect of this employment response on tax revenue depends on the tax payment that these individuals pay when entering the labor force and which is lost when they stay out.

At the intensive margin, there is a mechanical effect, a behavioral response from an income effect and a behavioral response from a substitution effect. The mechanical effect is that individuals with an income larger than y now pay additional taxes. This yields a revenue gain that is proportional to the mass of these people, $E_\Gamma\left[1 - F(\omega_0(y,\gamma) \mid \gamma)\right]$. With income effects, these people also seek to make up for the fact that the tax reform makes them poorer and they respond with an increase of their earnings. Together the mechanical and the income effect amount to $E_\Gamma\left[(1-I(\omega_0(y,\gamma)\mid\gamma))(1-F(\omega_0(y,\gamma)\mid\gamma))\right]$. The substitution effect is relevant for people with an income of y. They have less of an incentive to exert productive effort when the marginal tax rate for incomes close to y is increased. This is captured by the expression

$$T_0'(y)\;E_\Gamma\left[f(\omega_0(y,\gamma)\mid\gamma)\frac{y^*_{0\tau}(\omega_0(y,\gamma),\gamma)}{y^*_{0\omega}(\omega_0(y,\gamma),\gamma)}\right]\;.$$

For the special case of a utility function that is quasi-linear in consumption and has iso-elastic effort costs, this term can be written as

$$\frac{T_0'(y)}{1-T_0'(y)}\left(1+\frac{1}{\varepsilon}\right)E_\Gamma\left[f(\omega_0(y,\gamma)\mid\gamma)\right]\;,$$

where ε is the elasticity of earnings with respect to the net of tax rate.

The multiplier $\mathcal{M}^0$ mitigates both the intensive-margin and the extensive-margin effect. Its first component

$$E_\Gamma\Big[(1-F(\hat\omega_0(\gamma)\mid\gamma))\;I(\hat\omega_0(\gamma)\mid\gamma)\Big]\;,$$

is, again, reflecting income effects. With income effects, a tax reform that yields a revenue gain and hence increased transfers has a negative effect on the earnings labor of market participants. The increased transfers also play a role at the extensive margin. The sign of the total effect depends on the sign of $(P_0^{no}(\gamma)-P_0(\gamma))$. If $(P_0^{no}(\gamma)-P_0(\gamma))>0$, the increased transfers tend to make unemployment more attractive and this implies a loss of tax revenue.

D.1 Proof of Proposition D.1

The change in tax revenue associated with a one-bracket reform (τ,ℓ,y_a) can be written as

$$R^s(\tau,\ell,y_a)=\mathbb{E}_\Gamma\left[\mathcal{R}^s(\tau,\ell,y_a\mid\gamma)\right]\;,$$

where $\mathcal{R}^s(\tau,\ell,y_a\mid\gamma)$ is the change in tax revenue due to individuals with a given fixed cost type γ, and $\mathbb{E}_\Gamma$ is an expectations operator that indicates the computation of a population average using the marginal distribution F_Γ of fixed costs. Also, note that

$$\begin{aligned}\mathcal{R}^s(\tau,\ell,y_a\mid\gamma) &= \mathbb{E}_\Omega\left[T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma)-T_0(y_0^*(\omega,\gamma))\mid\gamma\right]\\ &= \textstyle\int_{\underline{\omega}}^{\overline{\omega}}\{T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma)-T_0(y_0^*(\omega,\gamma))\}f_\Omega(\omega\mid\gamma)\;d\omega\;,\end{aligned}$$

where $f_\Omega(\cdot\mid\gamma)$ is the density of the conditional distribution of ω for given γ. The change in revenue associated with a marginal change of τ can be written as

$$R^s_\tau(\tau,\ell,y_a)=\mathbb{E}_\Gamma\left[\mathcal{R}^s_\tau(\tau,\ell,y_a\mid\gamma)\right]\;,\tag{D.4}$$

where

$$\begin{aligned}\mathcal{R}^s_\tau(\tau,\ell,y_a\mid\gamma) &= \tfrac{d}{d\,\tau}\textstyle\int_{\underline{\omega}}^{\overline{\omega}}T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma)f_\Omega(\omega\mid\gamma)\;d\omega\\ &= \tfrac{d}{d\,\tau}\textstyle\int_{\hat\omega(\tau,\gamma)}^{\overline{\omega}}T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma)f_\Omega(\omega\mid\gamma)\;d\omega\;.\end{aligned}$$

By standard arguments,

$$\begin{aligned}\mathcal{R}^s_\tau(\tau,\ell,y_a \mid \gamma) &= \tfrac{d}{d\,\tau}\textstyle\int_{\hat\omega(\tau,\gamma)}^{\overline{\omega}} T_1(y^*(R^s(\cdot),\tau,h,\omega,\delta) f_\Omega(\omega \mid \gamma)\, d\omega \\ &= -T_1(y^*(R^s(\cdot),\tau,h,\hat\omega(\cdot),\gamma)) f(\hat\omega(\cdot) \mid \gamma)\, \hat\omega_\tau(\cdot) \\ &\quad + \textstyle\int_{\hat\omega(\tau,\gamma)}^{\overline{\omega}} \tfrac{d}{d\tau}\, T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma) f_\Omega(\omega \mid \gamma)\, d\omega\ ,\end{aligned} \tag{D.5}$$

i.e. the change of tax revenues can be decomposed into a change that comes from extensive margin responses and a change that comes from intensive margin responses to the tax reform.

A small change of the marginal tax rate. Equations (D.4), (D.5) and (D.1) imply that

$$R^s_\tau(0,\ell,y_a) = \mathbb{E}_\Gamma\left[\mathcal{R}^s_\tau(0,\ell,y_a \mid \gamma)\right]\ , \tag{D.6}$$

where

$$\begin{aligned}\mathcal{R}^s_\tau(0,\ell,y_a \mid \gamma) &= -T_0(y_0^*(\hat\omega_0(\gamma),\gamma)) f(\hat\omega_0(\gamma) \mid \gamma)\, \hat\omega_{0\tau}(\cdot) \\ &\quad + \textstyle\int_{\hat\omega_0(\gamma)}^{\overline{\omega}} \left(\tfrac{d}{d\tau}\, T_1(y^*(R^s(\cdot),\tau,h,\omega,\gamma)\right)_{|\tau=0} f_\Omega(\omega \mid \gamma)\, d\omega\ ,\end{aligned} \tag{D.7}$$

and

$$\hat\omega_{0\tau} = \frac{u^{no}_{0c}(\cdot)}{u_{0\omega}(\cdot)} R^s_\tau(0,\ell,y_a) - \frac{u_{0c}(\cdot)}{u_{0\omega}(\cdot)} \left(R^s_\tau(0,\ell,y_a) - h(y_0^*(\hat\omega_0(\gamma),\gamma))\right)\ . \tag{D.8}$$

By the arguments in Bierbrauer and Boyer (2018),

$$\begin{aligned}&\textstyle\int_{\hat\omega_0(\gamma)}^{\overline{\omega}} \left(\tfrac{d}{d\tau}\, T_1(y^*(R^s(\cdot),\tau,\omega,\delta)\right)_{|\tau=0} f_\Omega(\omega \mid \gamma)\, d\omega \\ &= R^s_\tau(0,\ell,y_a)\,\left(1 - F_\Omega(\hat\omega_0(\gamma) \mid \gamma)\right)\, I(\hat\omega_0(\gamma) \mid \gamma) + \mathcal{J}(\ell,y_a \mid \gamma)\end{aligned} \tag{D.9}$$

where

$$I(\hat\omega_0(\gamma) \mid \gamma) = E_\Omega\left[T_0'(y_0^*(s,\gamma))\; y_{0e}^*(s,\gamma) \mid s \geq \hat\omega_0(\gamma),\gamma\right]\ ,$$

as defined in the body of the text, and

$$\begin{aligned}\mathcal{J}(\ell,y_a \mid \gamma) &= \textstyle\int_{\omega_0(y_a|\gamma)}^{\omega_0(y_a+\ell\,|\gamma)} \left(T_0'(y_0^*(\omega,\gamma))\; y_{0\tau}^*(\omega,\gamma) + y_0^*(\omega,\gamma) - y_a\right) f(\omega \mid \gamma) d\omega \\ &\quad + \ell\,\left(1 - F(\omega_0(y_a+\ell \mid \gamma)) - \textstyle\int_{\omega_0(y_a+\ell\,|\gamma)}^{\overline{\omega}} T_0'(y_0^*(\omega,\gamma))\; y_{0e}^*(\omega,\gamma) f(\omega \mid \gamma) d\omega\right)\end{aligned}$$

and $\omega_0(y_a \mid \gamma)$ and $\omega_0(y_a + \ell\ \mid \gamma)$ are, respectively, the ω-types who choose income levels of y_a and $y_a + \ell$ in the status quo. Equations (D.2), (D.3), (D.7), (D.8) and (D.9) imply that

$$\begin{aligned}R^s_\tau(0,\ell,y_a \mid \gamma) &= R^s_\tau(0,\ell,y_a)\; M_0(\gamma) \\ &\quad - T_0(y_0^*(\hat\omega_0(\gamma),\gamma))\; P_0(\gamma)\; h(y_0^*(\hat\omega_0(\gamma),\gamma)) \\ &\quad + \mathcal{J}(\ell,y_a \mid \gamma)\end{aligned} \tag{D.10}$$

where the multiplier $M_0(\gamma)$ is given by

$$M^0(\gamma) := (1 - F(\hat{\omega}_0(\gamma)) \mid \gamma)\ I(\hat{\omega}_0(\gamma) \mid \gamma) - T_0(y_0^*(\hat{\omega}^0(\gamma), \gamma))\ \left(P_0^{no}(\gamma) - P_0(\gamma)\right)\ .$$

Equations (D.6) and (D.10) imply that

$$R_\tau^s(0, \ell, y_a) = (1 - E_\Gamma[M_0(\gamma)])^{-1}\Big(E_\Gamma[\mathcal{J}(\ell, y_a \mid \gamma)] - E_\Gamma[\mathcal{Y}(\ell, y_a \mid \gamma)]\Big)\ , \tag{D.11}$$

where

$$\mathcal{Y}(\ell, y_a \mid \gamma) = T_0(y_0^*(\hat{\omega}_0(\gamma), \gamma))\ \ P_0(\gamma)\ h(y_0^*(\hat{\omega}_0(\gamma), \gamma))\ .$$

If y_a is a very high, the tax reform affects only high incomes. Plausibly, y_a is then also above the income level that individuals at the extensive margin would consider, i.e. $y_a > \max_\gamma\ y_0^*(\hat{\omega}_0(\gamma), \gamma)$. In this case, $h(y_0^*(\hat{\omega}_0(\gamma), \gamma)) = 0$ and hence $\mathcal{Y}(\ell, y_a \mid \gamma) = 0$.

A small change of the marginal tax rate for a narrow bracket. We are interested in clarifying whether a reform that involves a small change of marginal tax rates for incomes in a narrow bracket can generate additional tax revenue. To this end, we provide a characterization of $R_{\tau l}^s(0, 0, y_a)$, i.e., of the cross-derivative of tax revenue with respect to the change of the marginal tax rate τ and the length of the bracket ℓ at the status quo. To understand the logic of this exercise, note that our previous derivations imply that $R_\tau^s(0, 0, y_a) = 0$, i.e. if the length of the bracket is zero, the change of the marginal tax rate applies to a null set of taxpayers. Consequently, there is no effect on tax revenue. If, however, $R_{\tau l}^s(0, 0, y_a) > 0$, then moving from $\ell = 0$ to some $\ell = \epsilon$ for $\epsilon > 0$ but small, implies that $R_\tau^s(0, \epsilon, y_a) > 0$, so that a small change of the marginal tax rate then has a positive effect on revenue. More formally, if $R_{\tau l}^s(0, 0, y_a) > 0$ there exist $\delta > 0$ and $\epsilon > 0$ so that $R^s(\delta, \epsilon, y_a) > 0$.

It follows from equation (D.11) that

$$R_{\tau l}^s(0, 0, y_a) = (1 - E_\Gamma[M^0(\gamma)])^{-1}\Big(E_\Gamma[\mathcal{J}_l(0, y_a \mid \gamma)] - E_\Gamma[\mathcal{Y}_l(0, y_a \mid \gamma)]\Big)\ , \tag{D.12}$$

where, by the arguments in Bierbrauer and Boyer (2018),

$$\begin{aligned} E_\Gamma[\mathcal{J}_l(0, y_a \mid \gamma)] \quad &= \quad T_0'(y_a) E_\Gamma\left[f(\omega_0(y_a, \gamma) \mid \gamma)\frac{y_{0\tau}^*(\omega_0(y_a, \gamma), \gamma)}{y_{0\omega}^*(\omega_0(y_a, \gamma), \gamma)}\right] \\ &\quad + E_\Gamma\left[(1 - I(\omega_0(y_a, \gamma) \mid \gamma))(1 - F(\omega_0(y_a, \gamma) \mid \gamma))\right]\ . \end{aligned} \tag{D.13}$$

We now work towards a characterization of $\mathcal{Y}_l(0, y_a) := E_\Gamma[\mathcal{Y}_l(0, y_a \mid \gamma)]$. Suppose that $y_a \leq \max_\gamma\ y_0^*(\hat{\omega}_0(\gamma), \gamma)$ and denote by $\gamma_0^*(y_a)$ and $\gamma_0^*(y_a + \ell)$, respectively, the types at the extensive margin who earn, respectively, incomes of y_a and $y_b = y_a + \ell$, i.e.

$$y_0^*(\hat{\omega}_0(\gamma_0^*(y_a)), \gamma_0^*(y_a)) = y_a\ , \quad \text{and} \quad y_0^*(\hat{\omega}_0(\gamma_0^*(y_a + \ell)), \gamma_0^*(y_a + \ell)) = y_a + \ell\ .$$

Armed with this notation, we can write

$$\begin{aligned} E_\Gamma[\mathcal{Y}_l(\ell, y_a \mid \gamma)] \quad &= \quad \textstyle\int_{\gamma_0^*(y_a)}^{\gamma_0^*(y_a+\ell)} T_0(\hat{y}_0^*(\gamma))\ P_0(\gamma)\ \ (\hat{y}_0^*(\gamma) - y_a)\ \ f_\Gamma(\gamma)\ d\gamma \\ &\qquad + \ell\ \textstyle\int_{\gamma_0^*(y_a+\ell)}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma))\ P_0(\gamma)\ \ f_\Gamma(\gamma)\ d\gamma\ , \end{aligned}$$

where $\hat{y}_0^*(\gamma)$ is a shorthand for $y_0^*(\hat{\omega}_0(\gamma), \gamma)$. Straightforward computations yield

$$\mathcal{Y}_\ell(0, y_a) = \int_{\gamma_0^*(y_a)}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma))\ P_0(\gamma) f_\Gamma(\gamma) d\gamma\ .$$

Now, to accommodate both the case $y_a \leq \max_\gamma\ \hat{y}_0^*(\gamma)$ and the case $y_a > \max_\gamma\ \hat{y}_0^*(\gamma)$, we will write henceforth

$$\mathcal{Y}_\ell(0, y_a) = \int_{\underline{\gamma}}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma))\ P_0(\gamma)\ \mathbf{1}(\hat{y}_0^*(\gamma) \geq y_a)\ f_\Gamma(\gamma) d\gamma\ .$$

Proposition D.1 as stated in the text now follows upon adopting the more concise notation

$$\mathcal{I}(y) := \mathcal{J}_\ell(0, y) \quad \text{and} \quad \mathcal{X}(y) := \mathcal{Y}_\ell(0, y)\ ,$$

for any earnings level y.

D.2 Implications of Proposition D.1

D.2.1 Diamond's model

With quasi-linear in consumption preferences and without fixed costs of labor market participation, we have

$$\mathcal{M}^0 = 0\ , \quad \text{and} \quad \mathcal{X}(y) = 0\ ,$$

and

$$\mathcal{I}(y) = T_0'(y)\ \ f(\omega_0(y))\ \frac{y_{0\tau}^*(\omega_0(y))}{y_{0\omega}^*(\omega_0(y))} + 1 - F(\omega_0(y))\ , \tag{D.14}$$

If effort costs are, moreover, iso-elastic, this can be rewritten as

$$\mathcal{I}(y) = -\frac{T_0'(y)}{1 - T_0'(y)}\ \ f(\omega_0(y))\ \ \left(1 + \frac{1}{\varepsilon}\right)^{-1} \omega_0(y) + 1 - F(\omega_0(y))\ . \tag{D.15}$$

Getting from equation (D.14) to equation (D.15) requires to invoke the first order condition characterizing $y_0^*(\omega)$ for the purpose of deriving comparative statics results that yield a characterization of $y_{0\tau}^*(\omega_0(y))$ and $y_{0\omega}^*(\omega_0(y))$. Thus,

$$R_{\tau\ell}^s(0, 0, y) = -\frac{T_0'(y)}{1 - T_0'(y)}\ \ f(\omega_0(y))\ \ \left(1 + \frac{1}{\varepsilon}\right)^{-1} \omega_0(y) + 1 - F(\omega_0(y))\ ,$$

which is equation (11) in the main text. Equation (12) can be derived from (D.14) using that

$$F_y(y) = F(\omega_0(y)) \quad \text{and} \quad f_y(y) = f(\omega_0(y))\ \omega_0'(y) = f(\omega_0(y))\ \frac{1}{y_{0\omega}^*(\omega_0(y))}\ .$$

and that, for any y in $y_0(\Omega)$, $\varepsilon = -\frac{y}{1 - T_0'(y)}\ y_{0\tau}^*(\cdot)$.

D.2.2 Fixed costs as an extension of Diamond's model

We now consider an extension of Diamond (1998) that includes fixed costs of labor market participation. Preferences are now given by

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

Again, the absence of income effect implies that $\mathcal{M}^0 = 0$. We can, once more, rewrite $\mathcal{I}(y)$ using the distribution of incomes F_y so that

$$F_y(y) = E_\Gamma \left[F_\Omega(\omega_0(y, \gamma) \mid \gamma) \right] \, .$$

This yields

$$\mathcal{I}(y) = -\frac{T_0'(y)}{1 - T_0'(y)} \, \varepsilon \, y f_y(y) \; + 1 - F_y(y) \, .$$

While this expression looks exactly as in Diamond's model, the distribution of incomes F_y is now shaped by the joint distribution of fixed and variable costs. We also rewrite

$$\mathcal{X}(y) = \int_{\underline{\gamma}}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma)) \; P_0(\gamma) \; \mathbf{1}(\hat{y}_0^*(\gamma) \geq y) \; f_\Gamma(\gamma) \; d\gamma \, .$$

in a way that is more handy in the context of our application: We first note that, with quasi-linear in consumption preferences, $P_0(\gamma) = f_\Omega(\hat{\omega}_0(\gamma) \mid \gamma)$, and therefore

$$\mathcal{X}(y) = \int_{\underline{\gamma}}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma)) \; \mathbf{1}(\hat{y}_0^*(\gamma) \geq y) \; f_\Omega(\hat{\omega}_0(\gamma) \mid \gamma) \; f_\Gamma(\gamma) \; d\gamma \, .$$

We then note that $\hat{y}_0^*$ is an increasing function.[38] Thus, if we denote by $\gamma_0(y)$ be value of γ for which $\hat{y}_0^*(\gamma) = y$, then we can write $\mathcal{X}(y)$ as

$$\mathcal{X}(y) = \int_{\gamma_0(y)}^{\overline{\gamma}} T_0(\hat{y}_0^*(\gamma)) \; f_\Omega(\hat{\omega}_0(\gamma) \mid \gamma) \;\; f_\Gamma(\gamma) d\gamma \, .$$

We seek an interpretation of $\mathcal{X}(y)$ in terms of extensive margin elasticities. Therefore, consider the following thought experiment: For pre-tax incomes in an interval $[y_0, y]$, after-tax incomes are marginally decreased. Overall employment is then reduced by

$$L_E(y_0, y) = \int_{\gamma_0(y_0)}^{\gamma_0(y)} \frac{f_\Omega(\hat{\omega}_0(\gamma) \mid \gamma)}{u_\omega(\;)} \; f_\Gamma(\gamma) \; d\gamma \, .$$

Denote the derivative of $L_E(y_0, y)$ with respect to y by $l_E(y)$ and apply Leibnitz rule to derive

$$l_E(y) = f_\Omega(\hat{\omega}_0(\gamma_0(y)) \mid \gamma_0(y)) \; u_\omega(\;)^{-1} \; f_\Gamma(\gamma_0(y)) \; \gamma_0'(y) \, .$$

[38] This follows from the following observations. First, recall that, by definition, $\hat{y}_0^*(\gamma) = y_0^*(\hat{\omega}_0(\gamma), \gamma)$. Second, with quasi-linear in consumption preferences, the function $y_0^* : (\omega, \gamma) \mapsto y_0^*(\omega, \gamma)$ is increasing in the first argument and constant in the second argument. Third, $\hat{\omega}_0 : \gamma \mapsto \hat{\omega}_0(\gamma)$ is increasing, as higher rents from labor market participation are needed to offset larger fixed costs.

Using the fundamental theorem of calculus, we can now also write

$$L_E(y_0, y) = \int_{y_0}^{y} l_E(y')\, dy' \ .$$

Thus, $l_E(y)$ is a marginal effect. One that measures the mass of people flowing out of the pool of people who earn y and into the pool of people who earn zero, when consumption for people earning y is slightly reduced. We will now rewrite $\mathcal{X}(y)$ in a way that highlights the significance of this employment response. We perform a substitution using

$$y' = \hat{y}_0^*(\gamma) \quad \text{and} \quad dy' = \frac{\partial\, \hat{y}_0^*(\gamma)}{\partial\, \gamma}\, d\gamma \ .$$

Also, note that $\gamma_0 : y' \mapsto \gamma_0(y')$ is the inverse of the function $\hat{y}_0^*$, so that

$$\gamma_0'(y')\, dy' = d\gamma$$

Finally, note that $\hat{y}_0^*(\gamma_0(y')) = y'$. Thus,

$$\begin{aligned}
\mathcal{X}(y) &= \int_y^{\overline{y}_0} T_0(y')\, \tfrac{f_\Omega(\hat{\omega}_0(\gamma_0(y'))|\gamma_0(y'))}{u_\omega(\ )}\, f_\Gamma(\gamma_0(y'))\, \gamma_0'(y')\, dy' \\
&= \int_y^{\overline{y}_0} T_0(y')\, l_E(y')\, dy' \\
&= \int_y^{\overline{y}_0} \tfrac{T_0(y')}{y' - T_0(y')}\, \tfrac{l_E(y')}{f_y(y')}\, (y' - T_0(y'))\, f_y(y')\, dy' \\
&= \int_y^{\overline{y}_0} \tfrac{T_0(y')}{y' - T_0(y')}\, \pi_0(y')\, f_y(y')\, dy' \ .
\end{aligned}$$

where $\overline{y}_0$ is the highest level of y in the support of F_y. We interpret

$$\pi_0(y') := \frac{l_E(y')}{f_y(y')}\, \big(y' - T_0(y')\big)$$

as an extensive-margin elasticity, it relates a percentage change in the mass of people at pre-tax income y due to extensive-margin responses to a percentage change in their after-tax labor income $y - T_0(y)$. The literature typically refers to $\mathbb{E}_y\left[\pi_0(y')\right]$ as the participation elasticity.

Upon collecting terms, and upon assuming an unbounded distribution of income in the status quo, we obtain

$$R^s_{\tau l}(0, 0, y) = 1 - F_y(y) - \varepsilon\, y\, f_y(y)\, \frac{T_0'(y)}{1 - T_0'(y)} - \int_y^{\infty} f_y(y')\, \pi_0(y')\, \frac{T_0(y')}{y' - T_0(y')}\, dy' \ . \quad \text{(D.16)}$$

Note that (D.16) coincides with equation (13) in the body of the text.

E Empirical analysis of the 2018 tax-transfer system

There were many reforms of the tax and transfer system since the mid-1970s. These reforms reduced the inefficiencies documented above, but they did not entirely eliminate them. To make this point, we plot the function $y \mapsto \mathcal{R}^s_{sp}(y)$ for the 2018 tax and transfer system in California.

Data description. We take account of the federal income tax and various transfer and welfare programs: the federal Earned Income Tax Credit, the Child Tax Credit (CTC, a partly refundable tax credit introduced in 1998), SNAP and Temporary Assistance for Needy Families (TANF, the successor of AFDC). As of 2018, the maximum EITC amounts to $5,716$ USD for single taxpayers with two children. It is phased in at a rate of 0.4 for annual incomes below $14,290$ USD and phased out at a rate of 0.2106 for incomes between $18,660$ and $45,802$ USD. Table E.1 below presents a summary of the sources we use to compute the US tax and transfer system in 2018 and the income distribution of single parents.

Table E.1: Sources for US tax-transfer system and income data, 2018

Information	Sources
Income tax	Internal Revenue Service, "1040 Instructions, Tax Year 2018", 2019. Accessible at `https://www.irs.gov/pub/irs-dft/i1040gi--dft.pdf`.
EITC	Tax Policy Center, "Earned Income Tax Credit Parameters, 1975-2018", 2018. Accessible at `https://www.taxpolicycenter.org/file/178859`.
AFDC	CCWRO, "CalWORKs (also known as AFDC/TANF)", 2018. Accessible at `https://www.ccwro.org/advocateresources/public-assistance-table`.
	City and County of San Francisco Human Services Agency, "CalWORKs Eligibility Handbook", 2018. URL: `https://www.sfhsa.org/file/6406`.
SNAP	New Mexico Department of Human Services, "Income Eligibility Guidelines for SNAP and Financial Assistance", 2018. URL: `https://www.hsd.state.nm.us/uploads/FileLinks/26463f122f47474487faee4922e09ce8/ISD_017_Income_Eligibility_Guidelines_for_SNAP_and_Financial_Assistance__FFY18_1.pdf`.
Income data	Current Population Survey, Annual Social and Economic Supplement (March wave) 2019. Accessible at `https://cps.ipums.org/cps/`.

Empirical results. Figure E.1 shows the effective marginal tax $T_0'(y)$ and the average tax $\frac{T_0(y)}{y}$ that result from the interplay of these programs. In 2018, effective marginal tax rates for low incomes are much lower than in the 1970s. Most notably, they are negative for incomes below 2,500 USD and in a small income range between 17,000 and 19,000 USD.

As before, we estimate the income distribution among single parents based on data from the March 2019 wave of the CPS. The share of single parents with no income has gone down considerably since the mid-1970s. It now amounts to 15.2%, about half of the share in the 1970s. Around 54.1% of single parents have strictly positive incomes below 45,802 USD and are, therefore, eligible for the EITC. In our data, 81.5% of single parents were female.

For the behavioral responses to taxation, we stick to the benchmark case with an intensive-margin elasticity of 0.33 and participation elasticities that are decreasing from 0.6 to 0.35 for incomes above 50.000 USD.

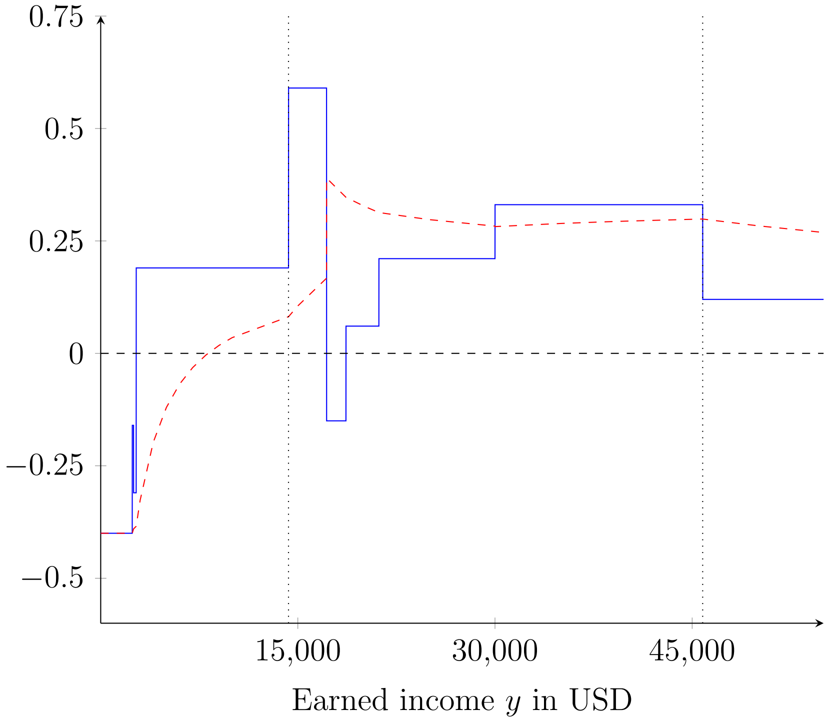


Figure E.1: US income tax and transfer schedule in 2018 for single parents

Notes: Figure E.1 shows the 2018 effective marginal tax $T_0'(y)$ (solid blue line) and the average tax $\frac{T_0(y)}{y}$ (dashed red line) for single parents as functions of earned income in 2018 USD. The 2018 EITC reduces effective marginal taxes in the phase-in range between 0 and 14,290 USD (first dotted line) and increases marginal taxes in the phase-out range between 18,660 and 45,802 USD (second dotted line).

Source: Authors' calculations (see Table E.1 for details).

Figure E.2 plots $y \mapsto \mathcal{R}_{sp}^s(y)$ for the US tax and transfer system in 2018. Remember that we identified both Pareto-improving one-bracket reforms and Pareto-improving two-bracket reforms in the mid-1970s: for incomes around 4,000 USD, $\mathcal{R}_{sp}^s(y)$ was negative and non-decreasing. The 2018 version of $y \mapsto \mathcal{R}_{sp}^s(y)$ is bounded from below by 0 and from above by 1. Hence, we no longer find Pareto improvements by means of one-bracket reforms. But there are still ranges where the function is non-decreasing, indicating that Pareto improvements by means of two-bracket reforms are still possible. The empirical evidence suggests that labor supply elasticities have decreased in magnitude over time. We therefore also considered an alternative calibration with smaller labor supply elasticities at both margins. Qualitatively, our results remain unchanged: the marginal revenue function satisfies the lower and upper boundary conditions, but violates the monotonicity condition. Hence, the 2018 tax and transfer system for single parents continues to be Pareto-inefficient, even though these inefficiencies appear less pronounced than in the 70s.

We abstain from measuring the size of the current inefficiency and relating it to its 1975 counterpart. Going into cardinal assessments of inefficiencies would lead us astray. That said, looking at the figures gives an intuitive sense that the contemporaneous inefficiencies are "smaller" than those in the 70s. We conjecture that a more systematic approach – e.g., one that invokes money-metric utility-based measures of the taxpayers' willingness to pay for a reform – would confirm this intuition.

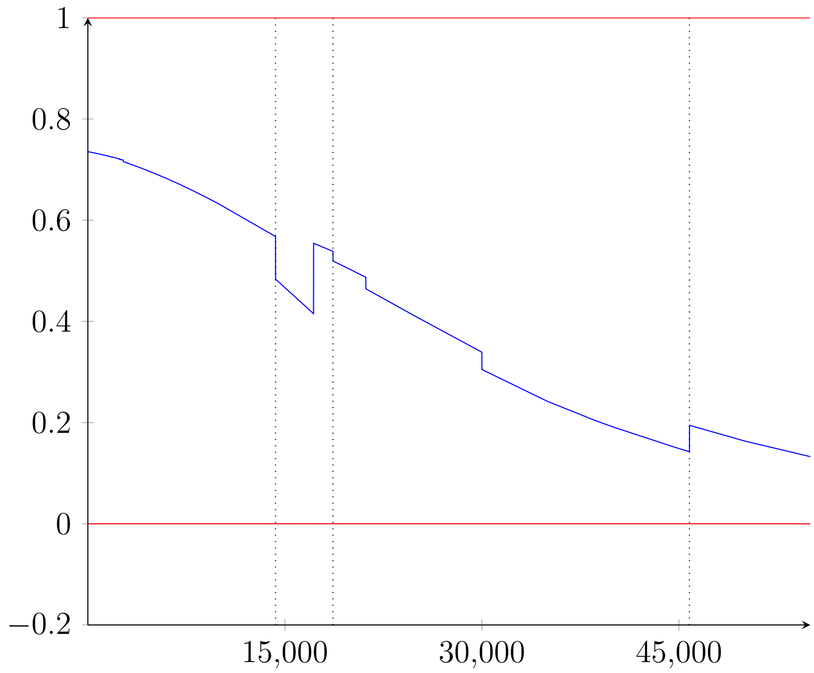


Figure E.2: Pareto (in)efficiency of the 2018 US tax and transfer system for single parents

Notes: Figure E.2 shows the marginal revenue functions for single parents $\mathcal{R}^s_{sp}(y)$ (blue, solid). The 2018 EITC reduces effective marginal taxes in the phase-in range between 0 and $14,290$ USD (first dotted line) and increases marginal taxes in the phase-out range between $18,660$ (second dotted line) and $45,802$ USD (third dotted line). The intensive-margin elasticity is fixed at 0.33 and the extensive-margin elasticity is assumed to be decreasing with income from 0.6 at very low incomes to 0.35 at incomes above 50,000 USD.

Source: Authors' calculations (see Table E.1 for details).