DISCUSSION PAPER SERIES

DP12328
(v. 2)

OPTIMAL INCOME TAXATION WITH UNEMPLOYMENT AND WAGE RESPONSES: A SUFFICIENT STATISTICS APPROACH

Etienne Lehmann, Kory Kroft, Kavan Kucko and Johannes Schmieder

PUBLIC ECONOMICS

ISSN 0265-8003

OPTIMAL INCOME TAXATION WITH UNEMPLOYMENT AND WAGE RESPONSES: A SUFFICIENT STATISTICS APPROACH

Etienne Lehmann, Kory Kroft, Kavan Kucko and Johannes Schmieder

Discussion Paper DP12328
First Published 22 September 2017
This Revision 11 February 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

OPTIMAL INCOME TAXATION WITH UNEMPLOYMENT AND WAGE RESPONSES: A SUFFICIENT STATISTICS APPROACH

Abstract

We derive a sufficient statistics optimal income tax formula in a general model that incorporâtes unemployment and endogenous wages, to study the shape of the tax and transfer system at the bottom of the income distribution. Key sufficient statistics are the macro employment response to taxation, the micro and macro participation response to taxation and the wage-moderating effect of tax progressivity. We empirically implement the tax formula by estimating the micro and macro elasticities using policy variation from the U.S. tax and transfer system. Our results suggest that the optimal tax more closely resembles a Negative Income Tax than an Earned Income Tax Credit relative to the case where unemployment and wage responses are not taken into account.

JEL Classification: N/A

Keywords: N/A

Etienne Lehmann - etienne.lehmann@gmail.com
CRED (TEPP) University Pantheon-Assas Paris II, IZA, CESifo and CEPR

Kory Kroft - kory.kroft@utoronto.ca
University of Toronto, NBER

Kavan Kucko - kavan.kucko@gmail.com
Cornerstone Research

Johannes Schmieder - johannes@bu.edu
Boston University, NBER and IZA

Acknowledgements
We would like to thank the Editor Dan Silverman, two anonymous referees, Felix Bierbrauer, Pierre Cahuc, David Card, Raj Chetty, Sebastian Findeisen, Peter Funk, Robert Gary-Bobo, Emmanuel Hansen, Bas Jacobs, Laurence Jacquet, Hilary Hoynes, Henrik Kleven, Bas Jacobs, Claus Thustrup Kreiner, Patrick Kline, Rafael Lalive, Kevin Lang, Thomas Le Barbanchon, Isabelle Mejean, Jean-Baptiste Michau, Austin Nichols, Matthew Notowidigdo, Claudia Olivetti, Daniele Paserman, Thomas Piketty, Jukka Pirttilä , Julien Prat, Jesse Rothstein, Dominik Sacks, Emmanuel Saez, Stefanie Stantcheva, Aleh Tsyvinski, Bruno Van der Linden, Owen Zidar as well as seminar participants at Boston University, McGill University, University of Connecticut, University of Toronto, CRED-University Paris II, CREST, TEPP, University of Cologne, Queens University, THEMA Cergy-Pontoise, GATE Lyon II, SOLE Meeting 2013, the NTA Meeting 2014, Berkeley, Hamilton, Wilfried Laurier, CEPR, SED meeting 2015, NBER Summer Institute, University of Chile, Lille 1, Nancy, Lausanne, Mannheim, Ecole Polytechnique, AFSE 2016, PET 2016, IIPF 2016 and PSE for many helpful comments on this project. All errors are our own.

Optimal Income Taxation with Unemployment and Wage Responses: A Sufficient Statistics Approach *

Kory KROFT†
University of Toronto, NBER

Kavan KUCKO‡
Cornerstone Research

Etienne LEHMANN§
CRED(TEPP), Paris II

Johannes SCHMIEDER¶
Boston University, NBER and IZA

Abstract

We derive a sufficient statistics optimal income tax formula in a general model that incorporates unemployment and endogenous wages, to study the shape of the tax and transfer system at the bottom of the income distribution. Key sufficient statistics are the macro employment response to taxation, the micro and macro participation response to taxation and the wage-moderating effect of tax progressivity. We empirically implement the tax formula by estimating the micro and macro elasticities using policy variation from the U.S. tax and transfer system. Our results suggest that the optimal tax more closely resembles a Negative Income Tax than an Earned Income Tax Credit relative to the case where unemployment and wage responses are not taken into account.

*We would like to thank the Editor Dan Silverman, two anonymous referees, Felix Bierbrauer, Pierre Cahuc, David Card, Raj Chetty, Sebastian Findeisen, Peter Funk, Robert Gary-Bobo, Emmanuel Hansen, Bas Jacobs, Laurence Jacquet, Hilary Hoynes, Henrik Kleven, Bas Jacobs, Claus Thustrup Kreiner, Patrick Kline, Rafael Lalive, Kevin Lang, Thomas Le Barbanchon, Isabelle Mejean, Jean-Baptiste Michau, Austin Nichols, Matthew Notowidigdo, Claudia Olivetti, Daniele Paserman, Thomas Piketty, Jukka Pirttilä , Julien Prat, Jesse Rothstein, Dominik Sacks, Emmanuel Saez, Stefanie Stantcheva, Aleh Tsyvinski, Bruno Van der Linden, Owen Zidar as well as seminar participants at Boston University, McGill University, University of Connecticut, University of Toronto, CRED-University Paris II, CREST, TEPP, University of Cologne, Queens University, THEMA Cergy-Pontoise, GATE Lyon II, SOLE Meeting 2013, the NTA Meeting 2014, Berkeley, Hamilton, Wilfried Laurier, CEPR, SED meeting 2015, NBER Summer Institute, University of Chile, Lille 1, Nancy, Lausanne, Mannheim, Ecole Polytechnique, AFSE 2016, PET 2016, IIPF 2016 and PSE for many helpful comments on this project. All errors are our own.

†Email: kory.kroft@utoronto.ca. Address: 150 St. George Street, 354 Toronto, ON, Canada.

‡Email: kavan.kucko@gmail.com. Address: Boston University, Department of Economics, 270 Bay State Road, Boston, MA 02215, USA.

§Email: etienne.lehmann@gmail.com. Address: CRED Université Panthéon-Assas Paris II, 12 Place du Panthéon, 75 231 Paris Cedex, France. Etienne Lehmann is also research fellow at TEPP, IZA, CESifo, IDEP and CEPR. Financial support from Institut Universitaire de France and Center for Equitable Growth at University of California Berkeley are gratefully acknowledged.

¶Email: johannes@bu.edu. Address: Boston University, Department of Economics, 270 Bay State Road, Boston, MA 02215, USA.

A large literature has sought to characterize the optimal income tax when there is a desire for redistribution (Mirrlees, 1971, Diamond, 1998, Saez, 2001). Many of the papers in this literature typically assume that wages are exogenous and there is full employment. The objective of this paper is to theoretically and empirically investigate the optimal income tax in the presence of endogenous wages and involuntary unemployment.

On the theory side, we build on the discrete model of Saez (2002). Individuals choose an "occupation" corresponding to a specific pre-tax income level (hereafter "wage") and tax liability. The model includes non-participation as an outside option and thus allows for selection into and out of the labor force. Saez interprets non-participation as unemployment. According to this definition, all unemployment is voluntary. We extend the model by introducing endogenous wages and involuntary unemployment, where some participants fail in their search to find a job. Gross wages and employment probabilities are assumed to be endogenous to the tax policy and this endogeneity is captured by very flexible reduced-form functions. We do not impose any a priori restrictions on these reduced-form functions and thus are able to capture a broad set of micro-foundations of the labor market with general equilibrium effects. Therefore, we adopt a "sufficient statistics" approach to characterizing the optimal tax policy (Chetty, 2009).

The first contribution of this paper is to derive a sufficient statistics optimal tax formula that is valid in the presence of endogenous wages and (involuntary) unemployment. Our formula generalizes the formula in Saez (2002) that includes both extensive and intensive labor supply responses.[1] It is expressed in terms of social welfare weights, the income distribution, *macroeconomic employment* responses to taxation, and *microeconomic* and *macroeconomic participation* responses to taxation.[2] By "micro" responses, we refer to responses to taxation, holding wages and the unemployment rate in all occupations constant. "Macro" responses, on the other hand, allow both wages and conditional employment probabilities to adjust to a change in taxes.

There are two main differences between our optimal tax formula and the one of Saez (2002). First, optimal tax formulas take into account the revenue effects of a tax reform. The revenue effects include both the mechanical change in tax revenue holding constant employment, as well as behavioral responses due to incentive effects. In Saez (2002), behavioral responses are captured by labor supply responses which are the same as micro employment responses due to the full employment and fixed wage assumptions. By contrast, in our model, labor supply responses are not sufficient to characterize revenue effects, since taxes affect equilibrium wages and unemployment through movements along the labor demand curve or search frictions. This drives a wedge between

[1] This is Equation (11) in the Appendix of Saez (2002), which corresponds to the general case.

[2] For ease of exposition, we hereafter refer to microeconomic as "micro" and macroeconomic as "macro". Participation responses refer to changes in participations rates (i.e. the fraction of participants, employed or unemployed, in a given population) to tax reforms while employment responses refer to changes in employment rates (i.e. the fraction of employed in a given population) to tax reforms.

labor supply and employment responses and as a result, behavioral responses are captured by macro employment responses to taxation which are the sum of the direct labor supply effect plus any indirect effects on employment operating through spillovers.

Second, optimal tax formulas take into account how taxation affects the social objective. In Saez (2002), this is pinned down by the mechanical effect of taxes on consumption. This is for two reasons. First, the analysis is partial equilibrium and taxes are assumed to only affect individual taxpayers (as opposed to firms). Second, labor supply responses have only second-order effects on the social objective due to the envelope theorem. Intuitively, labor supply responses due to a small change in taxes have no first-order effect on individuals' expected utility at an optimum; since individual utility is unaffected, social welfare is unaffected. In our model, taxes additionally impact wages and unemployment through spillover effects that operate in the market. These spillover effects have a direct impact on expected utilities, beyond the mechanical consumption effect, as they affect both gross wages (and hence consumption) and the probability of being employed.

We use a simple insight to characterize the social welfare effect in terms of sufficient statistics. We note that a tax reform alters the expected utility of participating in one occupation relative to non-participation. As a result, the magnitude of the participation response (in some occupation) to a given tax reform is informative about the magnitude of the utility change (in that occupation) generated by the tax reform. A tax cut that triggers a large increase in participation reveals that individuals experienced a large increase in expected utility. We use this intuition to show that when the social objective depends on individuals' expected utilities, the mechanical social welfare effects needs to be inflated by the (matrix) ratio between macro and micro participation responses. This is because both the social objective and participation decisions depend on taxation only through the effect of taxes on expected utility.

We then apply our formula to investigate whether the existence of unemployment and the endogeneity of wages reinforces or lessens the desirability of an Earned Income Tax Credit policy (a larger transfer to the working poor than to the non-employed, hereafter EITC) relative to a Negative Income Tax policy (i.e. a larger transfer to the non-employed than to the working poor, hereafter NIT).[3] For this purpose, we consider two special cases of the general model.

The first special case we consider assumes that the tax liability in a given occupation does not affect gross wages or conditional employment probabilities in other occupations. It also assumes that labor supply responses are concentrated along the extensive margin so that individuals can only participate in a single occupation or not participate at all. This is what we refer to as the

[3] The literature sometimes characterizes an EITC as a tax system featuring a negative marginal tax at the bottom, as in Saez (2002). Jacquet et al. (2013) show that this is a very local result: with a continuous income model, employment taxes feature a discontinuity at zero income and positive marginal tax rates at all income levels. Thus, we prefer to designate a tax system as corresponding to an EITC when the *employment tax* is negative at the bottom. Note that Hansen (2018) obtained both negative marginal and employment tax at the bottom.

"no-cross effect" case. In this case, the optimal tax formula has a simple form that can be directly compared to the optimal tax formula in the pure extensive margin models of Diamond (1980), Saez (2002), Choné and Laroque (2005, 2011). In particular, an EITC is desirable whenever the welfare weight on the working poor is larger than the ratio of the micro to the macro participation elasticities, while the threshold is one with fixed wages and full employment.

The second special case we consider assumes that gross wages and conditional employment probabilities in a given occupation depend on the tax liability in that occupation, the tax liabilities in two adjacent occupations, and the welfare benefit. Participation in a given occupation depends on expected utility in that occupation, along with the expected utilities in the two adjacent occupations and the utility for non-participants. This generalizes the "mixed model" of Saez (2002) with both intensive and extensive responses. There are key several differences that emerge. First, macro intensive and extensive responses appear instead of micro responses. This is because behavioral responses are captured by macro effects. Second, the social welfare weights need to be inflated to take into account that the macro effects of a tax reform on welfare differ from the micro ones. Finally, there is a new term since a change in the marginal tax rate may affect employment through a wage-moderating employment-enhancing effect of tax progressivity. This force tends to increase optimal marginal tax rates as in models with search unemployment (see Hungerbühler et al. (2006), Lehmann et al. (2011) and Hummel (2018)).

The second contribution of this paper is to numerically implement our optimal tax formula. Following Saez (2002), we restrict ourselves to our extension of the "mixed model ". We implement the formula in part by calibrating some of the sufficient statistics and in part by estimating some of them. In particular, we focus on estimating the macro employment response to taxation and the micro and macro participation response to taxation. In order to identify the micro effects of taxation, we need a tax reform that holds wages and unemployment responses constant in a labor market. In other words, we require variation in tax liabilities across individuals *within* the same labor market. To identify macro responses to taxation, we require variation in tax liabilities *between* labor markets. Our empirical strategy focuses on single individuals in the U.S. and exploits policy variation in the tax-and-transfer system that operates over time, across states and across households, which gives us three sources of policy variation. To estimate the micro participation response, we rely on time and household variation by exploiting reforms to the federal EITC that differentially affected single individuals depending on their number of children. The key assumption is that spillover effects uniformly affect individuals, across different family structures (e.g., different number of children), operating in the same labor market. Thus, our difference-in-differences estimate will only reflect micro responses. For the macro employment and participation response, we rely on policy variation across states and over time. Our identifying assumption is that spillover effects to state-specific tax policy operate within states, but not across them. Thus, our difference-in-differences estimate will capture both micro responses and spillover effects. To obtain both the micro and macro estimates, we isolate purely exogenous variation in tax liabilities coming from policy reforms, following the simulated instrumental variables approach in Currie and Gruber (1996) and Gruber and Saez (2002). Our instrumental variables (IV) estimates show that the micro participation elasticity, for the full sample of single individuals, is 0.57. Our estimate of the macro participation and employment elasticity is 0.48 and 0.42, respectively.

As an illustration, we use our empirical estimates to implement our sufficient statistics formula and calibrate the optimal income tax. We demonstrate three key results. First, the ratio of macro to micro participation responses has a large impact on the optimal tax policy: smaller macro responses move the optimal schedule more towards an NIT-like tax schedule with a relatively larger lump sum payment to the non-employed combined with higher employment tax rates and marginal taxes. In particular, our empirical point estimates of the macro-micro participation ratio imply that the optimal tax schedule features a lump sum payment that is around 17% higher than if this channel is ignored. Second, the employment-enhancing effect of tax progressivity also has a large impact on the optimal tax schedule the larger this effect, the larger the lump-sum payment, while at the same time increasing the employment tax rates, as well as marginal tax rates at incomes levels where employment taxes are positive. Third, we use our, albeit somewhat noisy, empirical estimates of behavioral responses over the business cycle to show that during recessions, the optimal income tax at the bottom shifts more towards an NIT-like structure with larger lump sum transfers and higher marginal tax rates.

Our paper contributes to several strands of the literature. In the optimal tax literature, our paper relates to Stiglitz (1982), Saez (2004), Rothschild and Scheuer (2013) and Sachs et al. (2016) who solve for the optimal income tax with endogenous wages, but do not consider (involuntary) unemployment. Hungerbühler et al. (2006) and Golosov et al. (2013) derive the optimal income tax in models with search frictions and involuntary unemploymen, but with exogenous wages. We show in the on-line appendix that these models are special cases of our model.

Our framework also highlights the relevance of the micro-macro "wedge" as a corrective device in optimal tax models with spillovers. This insight was initially introduced in an Unemployment Insurance (UI) context by Landais et al. (2018b,a) who derive a generalization of the "Baily (1978)-Chetty (2006)" optimal UI formula when there are spillovers. We show in the on-line appendix how their model can be obtained as a particular case of an extension of our model which includes a tax on profits. While Landais et al. (2018b,a) focus on a dynamic model in a social insurance setting, we consider the problem of redistribution in a static setting. Thus, our paper should best be seen as being complementary to their paper.

On the empirical side, we contribute to the empirical literature that estimates micro and macro behavioral responses to taxation. Chetty et al. (2011a,b), Chetty (2012) and Jäntti et al. (2015) have all tried to reconcile why taxation has a larger impact on employment at the macro (e.g., country) level compared to micro labor supply responses estimated from quasi-experimental research designs based on tax reforms. These studies use the terms "micro" and "macro" to correspond to the source of variation used to estimate elasticities. Macro estimates are identified from cross-country variation in tax rates whereas micro estimates are identified from quasi-experimental studies. Our contribution to this literature is to use the same dataset and a consistent method to estimate both the micro and macro elasticities for the U.S. We estimate the micro elasticity using micro data and control for labor market fixed effects. For the macro elasticity, we pool the data to the market level and control for year and state fixed effects and rely on within-state differences over time in the tax-and-transfer system. Thus, differences between micro and macro estimates are less likely to be confounded by differences in sample and/or methodology.

The rest of the paper proceeds as follows. Section I develops our theoretical model. Section II contains details on Institutional background and describes our data and empirical results. Section III considers the policy implications of our theoretical and empirical findings. The last section concludes.

I The theoretical model

The size of the population is normalized to 1. There are $I+1$ "occupations", indexed by $i \in \{0,1,...,I\}$. An "occupation" in this model is best interpreted as a technical construct that is distinct from the word's conventional meaning. It can be viewed as a set of jobs that pay the same wage. In the model, higher occupations require higher skills and thus higher wages. Occupation 0 corresponds to non-employment. We denote the gross wage in occupation i (equivalently pre-tax earnings) by w_i, the net wage (or consumption) by c_i and the tax liability by $T_i = w_i - c_i$.[4] The timing of the model is as follows:

1. The government chooses the tax policy.
2. Each individual m chooses an occupation $i \in \{0,...,I\}$ to participate in. This is modeled through heterogeneous search costs in each occupation.
3. For each occupation $i \in \{1,...,I\}$, a fraction $p_i \in (0,1]$ of participants are employed, receive gross wage w_i, pay tax T_i and consume the after-tax wage $c_i = w_i - T_i$. The remaining fraction $1 - p_i$ of participants are unemployed.

Unlike Saez (2002), we make a distinction among the *non-employed* individuals between the *unemployed* who search for a job in a specific occupation and fail to find one and the *non-*

[4] The assumption of a finite number of occupations is made for tractability. A continuous wage distribution can be approximated by infinitely increasing the number I of occupations.

participants who choose not to search for a job. For each occupation $i \in \{1,...,I\}$, k_i denotes the number of participants, $p_i \in (0,1]$ denotes the fraction of them who find a job and work, hereafter the *conditional employment probability*, and $h_i = k_i\, p_i$ denotes the number of employed workers. The number of unemployed individuals in occupation i is $k_i - h_i = k_i(1-p_i)$ and the unemployment rate is $1-p_i$. The number of non-participants is k_0. The number of non-employed is $h_0 = k_0 + \sum_{i=1}^{I} k_i(1-p_i)$.

We assume that all participants in occupation i face the same probability p_i to be employed. This "uniform rationing" assumption is made for tractability, just as the assumption that all employed in a given occupation are paid the same wage w_i. All the non-employed, whether non-participants or unemployed, receive the same welfare benefit denoted b. Therefore, the policy choice of the government is represented by the vector $\mathbf{t} = (T_1,...,T_I,b)$.

The government faces the following budget constraint:

$$\sum_{i=1}^{I} T_i\, h_i = b\, h_0 + E \qquad \Leftrightarrow \qquad \sum_{i=1}^{I} (T_i + b)\, h_i = b + E \tag{1}$$

where $E \geq 0$ is an exogenous amount of public expenditures. One more employed worker in occupation i increases the government's revenues by the amount T_i of tax liability she pays, plus the amount of welfare benefit b she no longer receives, the sum of the two defining the *employment tax*.[5] The budget constraint states that the sum of employment tax liabilities $T_i + b$ collected on all employed workers in all occupations finances the public good plus a lump-sum rebate b over all individuals.

In the on-line appendix, we consider several extensions of the model. First, we introduce a continuous job search intensity and obtain the same optimal tax formula. Second, we consider an extension of the model where unemployed in occupation i receive a benefit b_i that may differ from the welfare benefit z given to non-participants. We show that the generalized formula depends on the same sufficient statistics. Finally, we consider an extension of the model with a partial tax on profits and show that the wage response to taxation is an additional sufficient statistic that one must account for in the optimal tax formiula.

I.1 Micro vs. Macro Responses

In the paper, we make a crucial distinction between macro and micro responses to taxes. On the one hand, we define *micro* responses to a tax change in the hypothetical case where tax changes do not affect gross wages $w_1,...,w_I$ or conditional employment probabilities $p_1,...,p_I$. This is, for

[5] The literature interprets this term as the *participation tax*, which is not strictly correct whenever unemployment is present.

instance, the case for tax reforms frequently considered in the micro-econometric literature using difference-in-difference identification strategies. For these research designs, indirect or spillover effects of reforms typically affect wages and conditional employment probabilities in treatment and control groups in a uniform way. As a result, these research designs typically identify the micro or partial equilibrium effect of reforms.

On the other hand, *macro* responses to tax policy $\mathbf{t}$ are defined to encapsulate the general equilibrium responses of wages and conditional employment probabilities to taxes. To describe the latter, rather than specify micro-foundations, we use reduced-form functions denoted $\mathscr{W}_i(\cdot)$, $\mathscr{C}_i(\cdot)$ and $\mathscr{P}_i(\cdot)$. In occupation i, the gross wage is given by $w_i = \mathscr{W}_i(\mathbf{t})$, the net wage is given by $c_i = \mathscr{C}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathscr{W}_i(\mathbf{t}) - T_i$ and the conditional employment probability is given by $p_i = \mathscr{P}_i(\mathbf{t})$. At this general stage, we only assume that these functions are differentiable, that $\mathscr{P}(\cdot)$ takes values in $(0,1]$ and that $0 < \mathscr{W}_1(\mathbf{t}) < ... < \mathscr{W}_I(\mathbf{t})$ and $b < \mathscr{C}_1(\mathbf{t}) < ... < \mathscr{C}_I(\mathbf{t})$ for all tax policies $\mathbf{t}$. The two latter assumptions ensure that occupations indexed with a higher i correspond to occupations with higher before-tax and after-tax earnings.[6]

I.2 Labor supply decisions

Let $u(\cdot)$ be the cardinal representation of the utility individuals derive from consumption. This function is assumed to be increasing and weakly concave. Individual m faces in addition a utility cost $\chi_i(m)$ for searching for a job in occupation i, with the normalization $\chi_0(m) = 0$. Heterogeneity in $\chi_i(m)$ accounts for difference in the taste of work, but also for heterogeneity in skills. For instance, if individual m does not have the required skill to work in occupation i, it becomes extremely costly for her to apply to these jobs, in which case we consider that $\chi_i(m) = +\infty$.[7]

Individual m has a utility level equal to $u(c_i) - \chi_i(m)$ if she finds a job in occupation i, $u(b) - \chi_i(m)$ if she is unemployed in occupation i, and $u(b)$ if she chooses not to search for a job. Let

$$\mathscr{U}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathscr{P}_i(\mathbf{t})\, u\left(\mathscr{C}_i(\mathbf{t})\right) + \left(1 - \mathscr{P}_i(\mathbf{t})\right) u(b) \tag{2}$$

denote the *gross* expected utility of searching for a job in occupation i, absent any participation cost, as a function of the tax policy $\mathbf{t}$, and let:

$$U_i \stackrel{\text{def}}{\equiv} p_i\, u(c_i) + (1 - p_i) u(b)$$

denote its realization at a particular point of the tax system. Let $U_0 = u(b)$ be the utility expected

[6] This assumption plays a similar role in our model to the single-crossing property in Mirrlees (1971)

[7] We can extend the model to a utility cost that is specific to employed individuals in each occupation. The only consequence is notational.

out of the labor force. Individual m expects $U_i - \chi_i(m)$ by searching for a job in occupation i. She chooses to search in occupation i if and only if $U_i - \chi_i(m) > U_j - \chi_j(m)$ for all $j \in \{0,...,I\} \setminus \{i\}$. Provided that the distribution of participation costs $(\chi_1,...,\chi_I)$ is smooth enough, one can write the number k_i of participants in each occupation as a function denoted $\hat{\mathscr{K}}_i(\cdot)$ of gross expected utilities, so that:

$$k_i = \mathscr{K}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \hat{\mathscr{K}}_i(\mathscr{U}_1(\mathbf{t}),...,\mathscr{U}_I(\mathbf{t}),u(b)) \tag{3}$$

In other words, the tax policy influences participation decisions only through the determination of gross expected utilities. Therefore, to compute the micro and macro responses of participation to taxation, one need first to compute the micro and macro responses of gross expected utilities to taxation. The micro response of expected utility in occupation i to taxation in occupation j is given by:

$$\left.\frac{\partial \mathscr{U}_i}{\partial T_j}\right|^{\text{Micro}} = -p_i\, u'(c_i)\, \mathbb{1}_{i=j} \tag{4}$$

while the macro one is given by:

$$\frac{\partial \mathscr{U}_i}{\partial T_j} = \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j}\frac{u(c_i)-u(b)}{p_i\, u'(c_i)}\right] p_i\, u'(c_i) \tag{5}$$

The micro and macro responses of expected utilities to taxation differ for two reasons. First, for micro responses, unlike for macro ones, gross wages are held constant. Therefore, if $i = j$, the tax change is passed through one for one to the worker and $\frac{\partial \mathscr{C}_i}{\partial T_j} = -1$. If $i \neq j$, the tax reform has no effect on before-tax and after-tax wages, so $\frac{\partial \mathscr{C}_i}{\partial T_j} = 0$. Conversely, for macro responses, tax adjustments may affect wages in a variety of ways so in general, $\frac{\partial \mathscr{C}_i}{\partial T_j} \neq -\mathbb{1}_{i=j}$. Second, at the micro level, the conditional employment probabilities are unaffected by a change in taxes. Conversely, at the macro level, responses of conditional employment probabilities have to be taken into account. They may be due to changes in labor supply or due to changes in vacancy creation by employers following gross wage responses. The matrix $\left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{\text{Micro}}$ of micro responses of expected utilities is diagonal while the matrix $\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}$ of macro responses may not be diagonal, in which case the tax liability in one occupation affects expected utility in another occupation.[8] We refer to the off-diagonal elements of the matrix of macro responses as "cross effects" throughout.

Applying the chain rule to (3) and using matrix notations, the macro and micro participation

[8]For any function f of $\mathbf{t} = (T_1,...,T_I,b)$, we denote $\frac{\mathbf{df}}{\mathbf{dT}}$ the square matrix of rank I whose term in row j and column i is $\frac{\partial f_i}{\partial T_j}$ for $i,j \in \{1,...,I\}$. Symmetrically, the corresponding matrix of micro responses is denoted $\left.\frac{\mathbf{df}}{\mathbf{dT}}\right|^{\text{Micro}}$. Finally, we denote $\frac{\mathbf{d}\hat{\mathscr{K}}}{\mathbf{dU}}$ the square matrix of rank I whose term in row j and column i is $\frac{\partial \hat{\mathscr{K}}_i}{\partial U_j}$ for $i,j \in \{1,...,I\}$. In particular, these matrices do not include partial derivatives with respect to b. "$\cdot$" denotes the matrix product.

responses to taxation are given by $\frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}} = \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\hat{\mathscr{K}}}{\mathbf{dU}}$ and $\left.\frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}}\right|^{\text{Micro}} = \left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{\text{Micro}} \cdot \frac{\mathbf{d}\hat{\mathscr{K}}}{\mathbf{dU}}$. This lead us to the following lemma:

Lemma 1. *If matrix $\frac{\mathbf{d}\hat{\mathscr{K}}}{\mathbf{dU}}$ is invertible, the matrix ratio of macro to micro participation responses is equal to the matrix ratio of macro to micro responses of gross expected utilities, i.e.:*

$$\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{Micro} \right)^{-1} = \frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}}\right|^{Micro} \right)^{-1} \tag{6}$$

The effect of a government policy on individual welfare is summarized by how that policy affects her gross expected utility. Since participation responses depend only on expected utilities, it follows that one may use the macro and micro participation responses to characterize how the effects of a policy on welfare differ at the macro level compared to the micro one.

Lastly, given our definitions of labor force participation and job-finding probabilities, employment is given by:

$$h_i = \mathscr{H}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathscr{K}_i(\mathbf{t}) \ \mathscr{P}_i(\mathbf{t}) \tag{7}$$

I.3 Social objective

Choosing a social objective is a difficult issue as it relies on subjective value judgments. In this paper, we restrict the social objective to depend on the allocation only through the determination of expected utilities, so that the social objective is a function $\Omega\left(U_1,...,U_I,u(b)\right)$ of gross expected utilities, which is nondecreasing in each of its $I+1$ argument. This is for instance the case of any weighted utilitarian social objective of the form:

$$\Omega(U_1,...,U_I,u(b)) = \int \gamma(m) \left(\max_i \quad U_i - \chi_i(m) \right) \, d\mu(m)$$

where the weights $\gamma(m)$ are type-dependent and where $\mu(\cdot)$ is the distribution of individuals.

I.4 The optimal policy

The government chooses the tax policy $\mathbf{t} = (T_1,...,T_I,b)'$ to maximize the social objective $\Omega\left(U_1,...,U_I,u(b)\right)$ subject to the budget constraint (1). Let $\lambda > 0$ denote the Lagrange multiplier associated with the latter constraint. The Lagrangian of the government can be written using reduced-forms of the macro effects of taxes, independently of the micro-foundations underlying these reduced-forms:

$$\Lambda(\mathbf{t}) \stackrel{\text{def}}{\equiv} \sum_{i=1}^{I} (T_i + b)\ \mathscr{H}_i(\mathbf{t}) - b - E + \frac{1}{\lambda}\Omega\left(\mathscr{U}_1(\mathbf{t}), ..., \mathscr{U}_I(\mathbf{t}), u(b)\right) \tag{8}$$

Following Saez (2002), we define the marginal social welfare weight of workers in occupation $i \in \{1, ..., I\}$ as "the value (in terms of public funds) of giving an additional dollar to an individual in occupation i". Because we defined micro responses as those occurring for unchanged wages w_i and employment probabilities p_i, the mechanical increase in social welfare after a rise in consumption in occupation i is equivalent to the micro effect on social objective of a cut in tax liability in occupation i. We thus get:[9]

$$g_i \stackrel{\text{def}}{\equiv} -\frac{1}{\lambda\ h_i} \left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \frac{\partial \Omega}{\partial U_i} \qquad \Leftrightarrow \qquad \frac{1}{\lambda}\frac{\partial \Omega}{\partial U_i} = -\left(\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}}\right)^{-1} g_i\ h_i \tag{9}$$

The social weight g_i represents the social value in monetary terms of transferring an additional dollar to an individual working in occupation i ignoring the responses of wages and of the conditional employment probabilities. It is non-negative according to (4). Absent these responses, the government is indifferent between giving one more dollar to an individual employed in occupation i and g_i more dollars of public funds.

Using (9), the first-order condition with respect to the tax liability T_j in occupation j is:

$$0 = \underbrace{h_j}_{\text{Mechanical effect}} + \underbrace{\sum_{i=1}^{I} (T_i + b)\ \frac{\partial \mathscr{H}_i}{\partial T_j}}_{\text{Behavioral effects}} - \underbrace{\sum_{i=1}^{I} \frac{\partial \mathscr{U}_i}{\partial T_j}\left(\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}}\right)^{-1} g_i\ h_i}_{\text{Social Welfare effects}} \tag{10}$$

A unit increase in tax liability triggers the following effects:

1. **Mechanical effect**: Absent any behavioral response, a unit increase in T_j increases the government's resources by the mass h_j of employed individuals in occupation j.

2. **Behavioral effects**: A unit increase in T_j induces a change $\partial \mathscr{H}_i / \partial T_j$ in the level of employment in occupation i. This change incorporates general equilibrium (macro) responses. For each additional worker in occupation i, the government increases its resources by the employment tax $T_i + b$ that is equal to the additional tax received T_i plus the benefit b that is no longer paid.

3. **Social welfare effects**: A unit increase in T_j affects the expected utility in occupation i by $\partial \mathscr{U}_i / \partial T_j$. Multiplying by the rate $\frac{\partial \Omega}{\partial U_i}/\lambda$ at which each unit change in expected utility

[9] See Saez and Stantcheva (2016) when the social objective is not welfarist.

affects the social objective in monetary terms, we get that the social welfare effect of tax T_j in occupation i is: $\frac{\partial \mathscr{U}_i}{\partial T_j}\frac{1}{\lambda}\frac{\partial \Omega}{\partial U_i} = -\frac{\partial \mathscr{U}_i}{\partial T_j}\left(\frac{\partial \mathscr{U}_i}{\partial T_i}\Big|^{\text{Micro}}\right)^{-1} g_i\ h_i$. Note that because the social welfare function depends on expected utility U_i, the labor supply responses only modify the decisions of individuals that are initially indifferent between two occupations, and thus only have second-order effects on the social welfare objective, by the envelope theorem (Saez, 2001, 2002). Conversely, wage and unemployment responses are general equilibrium (macro) responses induced by the market instead of being directly triggered by individual choices. Thus, one must consider the macro responses of expected utility $\frac{\partial \mathscr{U}_i}{\partial T_j}$ instead of the micro ones $\frac{\partial \mathscr{U}_i}{\partial T_i}\Big|^{\text{Micro}}$. As the social weights are defined in terms of the micro effects of a tax change on the social objective, the usual social welfare terms $g_i\ h_i$ have to be inflated by the ratio between the macro and the micro responses of expected utility to taxation.

We now restate our optimal tax formula in terms of sufficient statistics. For this purpose, it is convenient to use matrix notation. Let $\mathbf{h} = (h_1, ..., h_I)'$ denote the vector of employment levels, let $\mathbf{g}\ \mathbf{h} = (g_1 h_1, ..., g_I h_I)'$ denote the vector of welfare weights times employment levels and let $(\mathbf{T} + \mathbf{b}) = (T_1 + b, ..., T_I + b)'$ denote the vector of employment taxes. We get:

Proposition 1. *If* $\frac{\mathbf{d}\hat{\mathscr{K}}}{\mathbf{dU}}$ *is invertible, the optimal tax system for occupations* $i = \{1, ..., I\}$ *solves:*

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathscr{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) - \frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}} \cdot \left(\frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}}\Big|^{Micro}\right)^{-1} \cdot (\mathbf{g}\ \mathbf{h}) \tag{11}$$

Proof: Equation (10) can be rewritten in matrix notations:

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathscr{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) - \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \left(\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\Big|^{\text{Micro}}\right)^{-1} \cdot (\mathbf{g}\ \mathbf{h}) \tag{12}$$

Using Equation (6) in Lemma 1 leads to (11). □

Due to the presence of wage and unemployment responses, our optimal tax formula differs from Equation (11) in Saez (2002) in two important ways. First, the behavioral effects depend on the macro responses of employment, not the micro labor supply responses. Second, as the welfare weights are computed from the micro effects of a tax reform on the social objective, while the macro responses are relevant to characterize the optimal tax, one needs a correcting procedure, which consists in multiplying the (vector of) welfare weights times employment levels by the (matrix) ratio of macro to micro responses of expected utilities, as in Equation (12). From (5), estimating this ratio requires estimating the macro responses of wages to taxation, which is especially difficult to identify in practice. However, as participation decisions depends on taxation only through expected utilities, according to Lemma 1, the matrix ratio of macro over micro responses of expected utility to taxation coincides with the matrix ratio of macro over micro participation responses, which can be identified more easily, as will be illustrated in the empirical Section II.[10]

I.5 The no-cross effect case

We now specialize our model in order to compare our results to the pure extensive model in Saez (2002). For this purpose, we consider the "no-cross effect" case where: $\partial \mathscr{W}_i / \partial T_j = \partial \mathscr{C}_i / \partial T_j = \partial \mathscr{P}_i / \partial T_j = \partial \hat{\mathscr{K}}_i / \partial U_j = 0$ for $i \neq j$ and $i \neq 0$. This means that labor demand only responds to tax liabilities in the occupation, but not in other occupations. It also implies that labor supply responses are concentrated along the extensive margin. In other words, individuals can move from non-employment to work (or vice-versa) in a single occupation, but cannot move between occupations in response to a tax change. The no-cross effect assumption implies from (5) that $\partial \mathscr{U}_i / \partial T_j = 0$, thereby: $\partial \mathscr{K}_i / \partial T_j = \partial \mathscr{H}_i / \partial T_j = 0$ for $i \neq j$, i.e. that the wage, the conditional employment probability, the employment level and the participation level in one occupation only depend on the welfare benefit b and on the tax liability in the same occupation, and not on tax liabilities in the other occupations. Therefore, the matrices of behavioral responses in (11) are all diagonal under the no cross-effect assumption.

Let the micro participation elasticity be defined as $\pi_j^m \overset{\text{def}}{\equiv} -\frac{c_j - b}{k_j} \frac{\partial \mathscr{K}_j}{\partial T_j}\Big|^{\text{Micro}}$, the macro participation elasticity as $\pi_j \overset{\text{def}}{\equiv} \equiv -\frac{c_j - b}{k_j} \frac{\partial \mathscr{K}_j}{\partial T_j}$ and the macro employment elasticity as $\eta_j \overset{\text{def}}{\equiv} -\frac{c_j - b}{h_j} \frac{\partial \mathscr{H}_j}{\partial T_j}$. From (7), the macro *employment* response η_j verifies $\eta_j = \frac{c_j - b}{p_j} \frac{\partial \mathscr{P}_j}{\partial T_j} + \pi_j$. It includes conditional employment responses $\frac{c_j - b}{p_j} \frac{\partial \mathscr{P}_j}{\partial T_j}$ in addition to the macro *participation* responses π_j. Equation (11) then simplifies to:

Proposition 2. *The optimal tax in occupation j in the no-cross effects case is:*

$$\frac{T_j + b}{c_j - b} = \frac{1 - \frac{\pi_j}{\pi_j^m} g_j}{\eta_j} \tag{13}$$

The no-cross effect environment is the simplest one to understand how the introduction of unemployment and wage responses modifies the optimal tax formula. In the pure extensive case without unemployment, Diamond (1980), Saez (2002) and Choné and Laroque (2005, 2011) showed that the optimal tax formula takes the form of an inversed elasticity rule $\frac{T_j + b}{c_j - b} = \frac{1 - g_j}{\eta_j}$. There are two key differences between the latter equation and Equation (13). First, the denominator in (13)

[10] We derive the optimal condition for the benefit level in Appendix A.1. We also show there that when income effects are assumed away, welfare weights sum to 1, i.e. $\sum_{i=0}^{I} g_i h_i = 1$.

corresponds to the macro employment elasticity. Second, Equation (13) inflates the social marginal welfare weight by the ratio of the macro to micro participation elasticity. As explained in Lemma 1, this is to account for the macro effects of taxation on the social objective, since the welfare weights are defined in terms of the micro expected utility responses. To understand why, consider a decrease in tax liability T_j in the no-cross effects case. This triggers a positive direct impact on social welfare $g_j h_j$, which is the only one at the micro level. Moreover, this decrease in tax liability typically induces a decrease in the gross wage when $\frac{\partial \mathscr{W}_j}{\partial T_j} > 0$, so the wage response attenuates the direct impact on social welfare. Finally, the decrease in tax liability also typically triggers a rise in job creation, whenever $\frac{\partial \mathscr{P}_j}{\partial T_j} < 0$, so the response of the conditional employment probability reinforces the direct impact on social welfare. The macro response of participation to taxation is therefore larger (smaller) than the micro one if the impact of the conditional employment response dominates (is dominated by) the impact of the wage response. Proposition 2 implies:

Corollary 1. *In the no-cross effect case, the optimal employment tax for the working poor is negative whenever* $g_1 > \frac{\pi_1^m}{\pi_1}$.

According to (13), a negative employment tax (EITC) becomes optimal whenever the social welfare weight is higher than the ratio of the micro over the macro participation elasticity, instead of one without unemployment and wage responses.

In principle, the macro participation elasticity can be either larger or lower than the micro participation elasticity. In the on-line appendix, we consider a search-and-matching model where for each occupation, unemployment arises because of matching frictions along (Mortensen and Pissarides, 1999), wages are set by proportional bargaining, the production function is linear and workers are risk neutral. This model is a structural version of our no-cross effect environment.[11] In such a model, a rise in tax liability is shared between the worker and the firm, so the after-tax wage decreases while the pre-tax wage increases. The latter reduction in turn induces firms to post less vacancies, which depresses labor demand and increases unemployment. The rise in pre-tax income tends to attenuate the effect of the tax increase on workers' consumption and leads to lower responses of expected utility at the macro level compared to the micro one. Conversely, the decrease in the employment probability tends to lead to lower responses of expected utility at the macro level compared to the micro one. We show that the micro and macro participation responses are equal under the Hosios (1990) condition where the workers bargaining power is equal to the elasticity of the matching function with respect to unemployment. Moreover, if the

[11] Under proportional bargaining (e.g. Jacquet et al. (2014)), the firm and the worker get a share of the surplus that does not depend on marginal tax rate, which is no longer the case under monopoly union (Hersoug, 1984), Nash bargaining (Pissarides, 1985, Lockwood and Manning, 1993, Hungerbühler et al., 2006) or competitive search models Lehmann et al. (2011) where a compensated increase in the marginal tax rate induces workers to concede wage moderation for higher employment.

worker's bargaining power is instead lower (larger) than the elasticity of the matching function the macro response is lower (larger) then micro one.

Estimating the ratio of the macro to micro participation elasticity is therefore sufficient to conclude whether unemployment and wages responses makes the EITC more or less optimal in the case where there are no cross effects.

I.6 The case with intensive and extensive responses

In this Section, we extend the optimal tax formula of Saez (2002, Equation (8)) with both intensive and extensive margins for the presence of wage and unemployment responses to taxation. For this purpose, we assume that wages $\mathscr{W}_i(\cdot)$ and employment probabilities $\mathscr{P}_i(\cdot)$ in a given occupation depends only on the tax liability T_i in that occupation, on tax liabilities in the two adjacent occupations T_{i+1} and T_{i-1} and on welfare benefit b. Moreover, participation in given occupation $\widehat{\mathscr{K}_i}(\cdot)$ depends only on expected utility U_i in that occupation, on expected utilities in the two adjacent occupations U_{i+1} and U_{i-1} and on utility out of the labor force $U_0 = u(b)$. Combining with Equations (2), (3) and (7), employment $\mathscr{H}_i(\cdot)$ in a given occupation therefore depends only on the tax liability T_i in that occupation, on tax liabilities in the two adjacent occupations T_{i+1} and T_{i-1} and on the welfare benefit b. We can therefore re-express employment in a given occupation as a function

$$\mathscr{H}_i(\mathbf{t}) \equiv \widetilde{\mathscr{H}_i}(T_i, T_{i+1} - T_i, T_i - T_{i-1}, b) \tag{14a}$$

of the tax liability T_i in that occupation, the two marginal tax rates $T_{i+1} - T_i$ and $T_i - T_{i-1}$ and the welfare benefit b. Let

$$\zeta_i \stackrel{\text{def}}{\equiv} -\frac{c_i - c_{i-1}}{h_i} \frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})} \tag{14b}$$

define the macro behavioral elasticity that captures intensive responses and - following Saez (2002) - is called the mobility elasticity. This can be related to the standard *intensive* elasticity of earnings with respect to marginal tax rates. In the competitive model of Saez (2002), this intensive margin elasticity is only due to individuals deciding to work in occupation $i-1$ instead of working in occupation i when the marginal tax rate $T_i - T_{i-1}$ increases, which implies $\frac{\partial \widetilde{\mathscr{H}_{i-1}}}{\partial (T_i - T_{i-1})} + \frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})} = 0$. Conversely, in models with wage and unemployment responses to taxation, a change in the marginal tax rate may trigger additional effects. For instance, a compensated increase in the marginal tax rate induces a wage moderation in many models of the labor market: it induces a substitution of employment for wage in monopoly union models (Hersoug, 1984) or in competitive search equilibrium models (Lehmann et al., 2011); it induces a substitution of profits for wage in wage bargaining models (Lockwood and Manning, 1993) or in search and matching models with Nash bargaining (Pissarides, 1985, 1998) or in efficiency wage models (Pisauro, 1991). Such wage moderating effects may in turn induce firms to hire more workers, thereby generating an "employment-enhancing effect of tax progressivity". We therefore introduce the following sufficient statistic to capture this effect:

$$\mu_i \stackrel{\text{def}}{\equiv} \frac{c_i - c_{i-1}}{h_i} \left(\frac{\partial \widetilde{\mathscr{H}_{i-1}}}{\partial (T_i - T_{i-1})} + \frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})} \right) \tag{14c}$$

which captures how much total employment in occupation i and $i-1$ change in response to a change in the marginal tax rate. Finally, changes in tax liabilities which are not associated with a change in marginal tax rates generate employment effects that are captured by:

$$\eta_i \stackrel{\text{def}}{\equiv} -\frac{c_i - b}{h_i} \frac{\partial \widetilde{\mathscr{H}_i}}{\partial T_i} \tag{14d}$$

Note that η_i captures both extensive responses and income responses.

Proposition 3. *When employment responses verify (14a), the optimal tax has to verify:*

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{\mu_i}{\zeta_i} \frac{T_{i-1} + b}{c_i - c_{i-1}} + \frac{1}{\zeta_i \, h_i} \sum_{j=i}^{I} \left(1 - \hat{g}_j - \frac{T_j + b}{c_j - b} \eta_j \right) h_j \tag{15}$$

where the inflated welfare weights $\hat{g}_1, ...\hat{g}_I$ are defined by the matrix identity:

$$(\hat{g}_1 \, h_1, ...\hat{g}_I \, h_I)^T = \frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}} \cdot \left(\left. \frac{\mathbf{d}\mathscr{K}}{\mathbf{dT}} \right|^{Micro} \right)^{-1} \cdot (\mathbf{g} \, \mathbf{h}) \tag{16}$$

To understand the economics behind the optimal tax formula (15) let us consider a small increase[12] dT in tax liabilities in occupations $j = i, i+1, ..., I$. We thus have: $dT = dT_i = dT_{i+1} = ... = dT_I$. This tax change raises $[h_i + h_{i+1} + ...h_I]dT$ additional taxes through the mechanical effect which are valued $[(1-\hat{g}_i)h_i + (1-\hat{g}_{i+1})h_{i+1} + ... + (1-\hat{g}_I)h_I]dT$ by the government.[13] The change in tax liabilities in occupations $j = i, i+1, ..., I$ also induces extensive and income responses that change the level of employment in occupations $j = i, i+1, ..., I$ by $dh_j = (\eta_j h_j / (c_j - b))dT$. These responses in turn modify the government's revenue by $(T_j + b)dh_j = \frac{T_j + b}{c_j - b} \eta_j h_j$. Finally, the marginal tax rate between occupation $i-1$ and i increases. This induces behavioral responses that can be decomposed in two terms. First, as in Saez (2002), a flow $(\zeta_i / (c_i - c_{i-1}))h_i$ of workers switch from occupation i to occupation $i-1$, which reduces government revenue by $\frac{T_i - T_{i-1}}{c_i - c_{i-1}} \zeta_i h_i$. Moreover, some additional jobs in occupation $i-1$ may be created through the employment en-

[12] The case $dT < 0$ where tax liabilities are decreased is obviously symmetric.

[13] See Appendix A.2 for a formal proof of why inflated welfare weights $\hat{g}_i$ show up instead of usual welfare weights g_i.

hancing effect, which implies a change in tax revenue equal to $\frac{T_{i-1}+b}{c_i-c_{i-1}}\mu_i h_i$.

There are three differences between our optimal tax formula (15) and the optimal tax formula of Saez (2002, Equation (8)) with both intensive and extensive responses. First, instead of the micro intensive and extensive elasticities, it is the macro mobility elasticity ζ_i and the macro tax liability elasticity η_i that show up in (15). This is because behavioral responses are not only generated by micro responses. Moreover, we do not rule out income effects. Second, the welfare weights need to be adjusted to take into account that the macro effects of a tax reform on welfare differ from the micro ones. This is the reason why it is the inflated welfare weights $\hat{g}_i$ that show up and not the usual welfare weights g_i. Last, a change in the marginal tax rate may not only generate a flow of employment between two consecutive occupations, but may also increase the overall employment level, for instance through a wage moderating - employment enhancing effect. In that case, one gets that $\mu_i > 0$. Moreover, if the optimal employment tax $T_i + b$ is positive, this employment-enhancing effect of tax progressivity is an additional force that tends to increase optimal marginal tax rates as in optimal tax models with search unemployment of Hungerbühler et al. (2006), Lehmann et al. (2011) and Hummel (2018).

II Estimating Sufficient Statistics

Our optimal tax formula (15) shows that the macro employment response to taxes, and the micro and macro participation responses are among the set of sufficient statistics. While a large number of studies have estimated micro responses, few have estimated macro responses and many do not distinguish between participation and employment responses. Our objective in this section is to estimate micro and macro extensive-margin responses for both employment and participation using a consistent methodology for the U.S. population.

We take a local labor market approach and define a geographic labor market by location (state), education and time (year-month). To identify macro responses, it is necessary to have exogenous variation in tax liabilities at the level of the local labor market. For micro estimates, we need exogenous variation between individuals within a given labor market.

II.1 Data

Current Population Survey (CPS)

We follow the large empirical literature on the effects of the Earned Income Tax Credit (EITC) and welfare reform in the U.S. and focus on low-income, single individuals throughout the last three decades. As a consequence of the gradual expansion of the EITC and the 1990's welfare reform, this group experienced substantial changes in participation and marginal tax rates differentially by number of children, within and across states.

Our analysis is based on data from the monthly outgoing rotation group (ORG) and the March annual data of the Current Population Survey (CPS). The March annual data spans the time period 1984-2011, while the ORG data (from IPUMS) spans 1994-2010. As our analysis sample, we select all single individuals age 18 to 55 who are not in the military or enrolled full time in school or college. Since there is insufficient tax variation for higher income individuals over our sample period, and since lower income individuals are more likely to be on the margin of the labor force, we further restrict the sample to individuals with education less than a bachelors degree. Our theory distinguishes between individuals who choose to participate in the labor force (and are employed or unemployed) and those individuals who are actually employed. We measure these labor market states using the standard International Labor Office (ILO) criteria. A person is classified as being in the labor force if she is either employed or unemployed (i.e., actively looking for a job during the reference week and was available for work) and employed if she has been working during the reference week (or been temporarily absent from a job).[14]

Panel A of Table 1 shows descriptive statistics for the demographic characteristics of single individuals in the March CPS for the full sample (Column 1) and broken down by educational attainment groups (Columns 2-4), pooling all years from 1984 to 2011.[15] The age range is pretty similar across the three education groups - less than high school, high school, and some college - but there are large differences in the distribution of number of children, with lower educated single individuals being more likely to be parents. This is likely due to our sample restriction to singles since higher educated parents are more likely to be married. Additionally, low educated individuals are more likely to be black or Hispanic than high educated ones. Panel B displays labor market variables by educational attainment. Lower educated individuals are much less likely to be in the labor force than higher educated ones and also experience higher unemployment rates.

Tax and Transfer Calculator

In order to estimate the employment and participation effects of taxes and transfers it is necessary to compute the budget sets that individuals face. For this purpose, we developed a calculator that computes taxes and transfers at (nominal) income levels for single individuals, depending on the number of children, state and year. We assume that an individual is filing as the head of the household and is claiming his/her children as dependents. To compute taxes (covering federal and state income taxes, including tax credits, as well as FICA liability), we rely on the NBER TAXSIM software. We assign taxes based on state of residence, as reported in the CPS, as well as number of

[14] For complete details on sample construction and variable definitions, please see the on-line appendix.

[15] We do not include the CPS ORG in this table since it spans different years, but when we compare sample means for the March CPS and ORG for the same period they are extremely close.

children, year, and income.[16] To compute transfers, in particular Aid to Families with Dependent Children (AFDC), Temporary Assistance for Needy Families (TANF) and Supplemental Nutrition Assistance Program (SNAP), we construct a benefit calculator based on rules published in the Welfare Rules Database, managed by the Urban Institute. This allows us to compute the benefits an individual is eligible for, as a function of number of children, state of residence, year and income. The shift from AFDC to TANF introduced a number of additional work and eligibility requirements for welfare recipients. For example, federal rules require a minimum number of TANF recipients to be employed and the lifetime duration of receiving TANF benefits is limited to a total of 5 years.[17] Rather than incorporate all of these policies explicitly into our empirical framework, we multiply benefits by gender specific recipiency rates constructed from the Survey of Income and Program Participation (SIPP). The new eligibility requirements are reflected in lower observed recipiency rates in our sample post-welfare reform.

We use our tax and transfer calculator to compute the incentive to work. Since we focus solely on the extensive margin in our analysis, we capture work incentives using just two measures, the transfer an individual receives when she has zero income and the tax and transfer level at the earnings level an individual obtains when working. A key difficulty is that earnings, and hence tax liabilities, are unobserved for non-employed individuals. Moreover, earnings for employed workers may be endogenous to the tax system. We proceed using two approaches. First, we impute an individual's tax liability following the approach taken in Eissa and Hoynes (2004) and Gelber and Mitchell (2012). We run separate regressions for each education group and year of log annual earnings for individuals on state fixed effects and control variables.[18] The control variables include a quadratic function of age, dummy variables for black and hispanic, and a categorical variable describing geographic location (i.e., urban versus rural). For each individual in our sample (both the non-employed and employed), we construct predicted earnings using the regression coefficients estimated from our model. This is for the purpose of obtaining a consistent specification.[19] We

[16] For an individual who resides and works in different states, the following rules apply. Generally an individual is required to pay income tax to his or her state of residence first. Then they must file as a non-resident in the state where they work, but get to take the amount of tax paid to the state of residence as a tax credit, and only pay the difference. If the amount of tax paid to the state of residence is greater than the tax bill for the work state, the individual doesn't pay anything to the work state, but still has to file. We don't take this into account in computing tax liabilities.

[17] In general, a state must have 50 percent of its single parent households and 90 percent of its dual parent households engaged in work-related activities (these include not only work but searching for work or training courses) for a minimum number of hours per week (30 hours per week or 20 hours if there is a young child). The 50 percent and 90 percent are calculated from a pool of "work-eligible individuals" which does not include single parents of children under the age of 1. States can obtain credits against the 50 and 90 percent rates for overall caseload reduction.

[18] For this exercise, we use earnings from the March CPS. To deal with misreporting we also drop observations where the implied hourly wage is less than one dollar or greater than one hundred dollars.

[19] As an alternative, we tried performing a Heckman selection correction to control for self-selection using the number of children and the presence of young children in the selection equation. However, we found that the pattern of results were not very well behaved. In particular, predicted earnings for high school dropouts seemed too high and earnings for higher education levels seemed unrealistically low relative to the raw differences earnings across

then use predicted earnings to impute an individual's tax liability using TAXSIM and the benefit calculator described above. In the on-line appendix, we present OLS regressions of participation and employment using this imputed tax liability.

One problem with this approach is that the demographic distribution itself, and therefore the imputed tax liabilities, might be endogenous to tax policy. For instance, more generous transfers to singles with kids, but not to individuals without children, may boost fertility and impact earnings. To address this concern, we also rely on a simulated instrument approach based on Currie and Gruber (1996).[20] This approach isolates policy variation in tax liabilities since it uses a fixed income and demographic distribution during the sample period.

Panel C of Table 1 shows the mean imputed real earnings for each education group averaged over the years and the corresponding tax and transfer levels depending on the number of children in the household. All numbers are reported in real 2010 U.S. dollars (USD). For high school dropouts, taxes (transfers) are strongly decreasing (increasing) in the number of children. This partly reflects the EITC which reduces tax liabilities according to the number of children in the household. The welfare benefit for households with no children is driven entirely by SNAP since these households are ineligible for AFDC/TANF. For bachelor degree holders (not shown), the range is very small and close to 0 since most are ineligible for these mean-tested benefits. Importantly, the reported welfare benefits do not incorporate recipiency rates which are much less than 100 percent during our sample period. The last four rows report recipiency rates, as estimated in the SIPP. Each individual in the CPS is assigned a recipiency rate that we calculate from the SIPP based on gender, education, income and year for individuals in the eligible income range. The table reports the average of the assigned recipiency rates separately for AFDC/TANF and food stamps, and also pre- and post-1996. We see that for high school dropouts, recipiency rates are roughly 50 percent for AFDC/TANF but fall to 20 percent post-1996. For food stamps, recipiency rates are much more comparable pre- and post-1996 and equal to roughly 40 percent.[21] These recipiency rates decrease with education which reflects diminishing eligibility as earnings increase.

II.2 Empirical Method

Let i denote households, e denote education level, n denote the number of children in the household, s denote state of residence and t denote time. Our econometric specification to estimate the micro participation response is given by:

education groups. This is likely due to the lack of a convincing instrument for working.

[20]Gruber and Saez (2002) use this approach to estimate taxable income elasticities; however, we are not aware of any papers that use this approach to estimate extensive margin labor supply responses.

[21]For AFDC/TANF, we calculate recipiency rates based on sample of single parents since singles with no children are not eligible for these programs.

$$k_{i,e,n,s,t} = T_{e,n,s,t}\beta + \delta_{e,s,t} + \delta_{e,n} + X_{i,e,n,s,t}\lambda + \varepsilon_{i,e,n,s,t} \tag{17}$$

The employment tax liability is defined as $T_{e,n,s,t}$ and it is the tax liability at an income level plus the transfers at 0 income. Implicit in this definition is the assumption that there are no income effects.[22] The term $\delta_{e,s,t}$ denotes labor market fixed effects interacted with education groups (one FE for each state-by-year-by-month-by-education cell), $\delta_{n,e}$ are number of children-by-education group fixed effects, $X_{i,e,n,s,t}$ are controls (demographic variables such as age, age-squared, race, and ethnicity all interacted with education groups).[23] The coefficient β captures the micro participation effect.

Next, to estimate macro participation responses, we aggregate the data to state-year-education averages to get:

$$k_{e,s,t} = T_{e,s,t}\gamma + \delta_{e,s} + \delta_{e,t} + X_{e,s,t}\lambda + \varepsilon_{e,s,t} \tag{18}$$

Here, $T_{e,s,t}$ is the macro employment tax, $\delta_{e,s}$ and $\delta_{e,t}$ are education-by-state and education-by year fixed effects, respectively and $X_{e,s,t}$ is a vector of controls, including region-specific linear time trends, demographic characteristics (cell averages of the micro controls) interacted with education. The macro effect, γ is defined as the change in individual participation probabilities if the tax liabilities for all individuals in a labor market increase by one dollar. Finally, to estimate macro employment responses, we simply replace participation k with employment h.

In our model, the behavioral responses are measured in terms of the number of individuals employed in and participating in the labor market. However, for an empirical specification that uses variation across individuals and labor markets, it makes little sense to use levels and assume constant effects (across labor markets). Instead, we will estimate the effect of taxes on employment and participation *rates*. Estimating the marginal effects of taxes on employment and participation rates furthermore has the important advantage that the estimates are easier to interpret and compare to the prior literature.

Identification

To identify the parameter β, we require that the micro tax liability $T_{e,n,s,t}$ is exogenous, conditional on labor market and education-by-number of children fixed effects and observables. Similarly, our identifying assumption for γ is that the macro tax liability $T_{e,s,t}$ is exogenous, conditional on education-by-state and education-by-year fixed effects and observables. Thus, two independent

[22] We tested whether the condition $\frac{\partial K_i}{\partial T_i} = \frac{\partial K_i}{\partial b}$ holds and found that the difference was very small and statistically insignificant. We therefore only report results under the no income effect assumption.

[23] Moffit (1998) argues that the literature features very heterogeneous marriage and fertility responses to taxes and transfers across studies, with a large number of studies finding no effect. As a result, he concludes that much more research remains to be done.

sources of exogenous variation in tax liabilities are needed. For the micro response β, we require variation in tax liabilities across individuals *within* the same labor market. For the macro response γ, we require variation in average tax liabilities *between* labor markets.

As described above, our strategy is to generate such variation using a simulated instrument approach. The policy variation in the micro tax liability is illustrated in Figure 1a). This figure plots the average value of the micro simulated tax liability, by year and number of children, relative to the value in 1984, for high school dropouts. One can see that there is substantial variation in taxes over time and this variation is very different across the number of children. Much of this is driven in large part by the EITC. In particular, the TRA86 reform can be clearly seen in 1986-1987, but is quite small relative to the expansions in the 1990s, which also introduced differential EITC levels for parents with one or two children. Finally in 2009, the EITC was expanded for parents with 3 children, as can be seen in the figure, and income taxes were cut for all family types. The identification strategy is similar to the one used by Eissa and Liebman (1996), Meyer and Rosenbaum (2001) and Gelber and Mitchell (2012).

The policy variation for the macro tax liability comes mainly from changes in state income taxes; in particular, the state-level EITCs and welfare benefits, which vary across states and over time. The large expansions of the federal EITC, that much of the literature has relied on, are not useful, since the change affected all states simultaneously and thus would be collinear with time trends. We illustrate this variation by plotting the macro simulated tax liability for high school dropouts for the largest 12 states in Figure 1b).

A potential concern with our identification strategy is that individuals might move to avoid taxes and/or receive higher benefits. However, several papers (e.g. Meyer, 2000, Kennan and Walker, 2010) suggest that this response is at best modest, particularly for the sample of low income individuals that are the focus of this study. Thus, while migration responses might be important in other contexts, we do not believe that our estimates will be confounded by them.

II.3 Empirical Results

For all of our empirical results, we report Instrumental Variables (IV) estimates from a Two-Stage Least Squares (2SLS) regression. Reported standard errors in all regressions are clustered on the state level. The notes of the tables contain exact details about the regression specification. All of the OLS results can be found in the on-line appendix. Note that in interpreting these results that the tax liabilities are in units of $1000.

The top panel of Table 2 shows the IV estimates for the micro participation (Column 1) and employment (Column 2) responses to taxes and transfers based on equation (17) above. The results indicate a clear negative and statistically significant participation effect of taxes, consistent with the prior literature. We find that a \$1000 increase in taxes leads to a 3.1 percentage point reduction in the participation probability which translates to an elasticity of -0.57.[24] We also see fairly similar micro responses for employment.[25]

Our elasticity estimates are well within the range of elasticities that is reported in the literature.[26] This is not that surprising since we use similar variation in taxes as the previous literature; in particular, variation driven by the EITC. One notable difference is that past studies typically control for state and year fixed effects, but not their interaction. This yields estimates that confound micro and macro responses (See Rothstein (2010) for a discussion of this). Nevertheless, most of the tax variation in these papers would also have come from across group variation within labor markets.

The macro participation and employment IV estimates are displayed in the second panel of Table 2. These correspond to empirical estimates from a macro-level (education-state-year cells) 2SLS regression of participation and employment rates on market-level tax liabilities, controlling for education-by-state and education-by-year fixed effects and percent black, percent Hispanic, average age, average age-squared, average number of children and their interactions with education and region-specific time trends. Since the number of observations is much smaller and since there is less variation in tax liabilities across labor markets, the coefficients are estimated less precisely. Nevertheless, there is some suggestive evidence that the macro participation and employment responses are smaller than the micro ones. We show in the online appendix (Proposition A.2) that such a finding is consistent with a matching model where the bargaining power is lower than the one prescribed by the Hosios condition.

Our results on micro and macro responses to taxation are generally consistent with the meta analysis conducted in Chetty et al. (2012) who report slightly larger estimates of the extensive steady-state elasticities based on micro evidence than macro evidence. It is worth noting that the macro-based studies cited in Chetty et al. (2012) are based on cross-country evidence that typically comes from a limited number of OECD countries. Nevertheless, it is reassuring to note that our results are similar, based on a panel data approach across all states, over time, in the U.S.

A concern with our macro estimates, which are identified by state-year variation in tax liabilities, is that they may be confounded by policy endogeneity. In particular, states may endogenously

[24] Following the theory, we take the marginal effect and multiply it by the ratio of the income gain from employment over the participation rate. These numbers are all reported in Table 2.

[25] The Online Appendix reports the OLS regression results. We see that the OLS participation responses are attenuated relative to our IV estimates. For the full sample, the micro participation elasticity is 0.09 and the macro participation elasticity is -0.8. The micro and macro employment responses are of a similar magnitude. This highlights the importance of instrumenting for the micro and macro tax liabilities. In general the OLS results are not very informative, for example there is a strong reverse causality issue where high participation rates will be associated with lower earnings (due to selection) and higher employment taxes. Isolating variation coming from tax policy changes is crucial in order to obtain meaningful results.

[26] Eissa et al. (2008) report a range of $(-0.35, -1.7)$ with a central elasticity of -0.7.

set taxes and welfare benefits based on prevailing local economic conditions. Our baseline estimates control for region-specific time trends which should partially address this issue. To further explore the robustness of our estimates, we consider several alternative specifications and report the results in Table 3. Table 3 provides a series of robustness tests. The first column reports our baseline estimates for comparison. In columns 2-4 we drop the region-specific time trends from the regressions and include alternative controls for pre-trends. Since the micro participation regressions control for year-by-state fixed effects, these are not affected (Panel A), but Panel B and C show that the macro responses are very robust to controlling for division-by-year fixed effects, region-by-year fixed effects and no controls for pre-trends. In column 5 we present our results dropping state taxes (state EITC and state income taxes) from our imputed tax liability and instrument, as those may be endogenous, as Hoynes and Patel (2015) have argued. While this slightly reduces the precision of our macro estimates, the results are qualitatively similar. Finally, column 6 controls for the state unemployment rate interacted with education as a proxy for the state specific economic environment and shows a very similar pattern. Overall, the robustness of our estimates suggest that policy endogeneity is not of first-order importance in our setting.[27]

Finally, Table 4 considers behavioral responses over the business cycle. In particular, this allows us to test whether spillovers are larger in recessions, as some recent research has found. We rely on several proxies for the business cycle: the 6-month change in the unemployment rate, the state unemployment rate and an indicator for whether the unemployment rate exceeds 9 percent. Across all specifications, we see that micro and macro participation and employment responses tend to be lower when the unemployment rate is relatively high. This is consistent with results in Schmieder et al. (2012) and Kroft and Notowidigdo (2016). There is also some suggestive evidence that the micro-macro participation gap increases in weak labor markets; for instance, for the 6-month change in unemployment specification, the gap is roughly 0.01 in weak labor markets but only 0.002 in strong labor markets. We emphasize however, that lack of precision limits any strong conclusion about how the gap varies over the cycle.

Overall, these results suggest that while micro labor supply responses are sizeable and in line with what the literature has found before, they may not always be good approximations for the macro employment responses. In particular our evidence broadly suggests that macro responses tend to be lower than micro responses. Although this is some of the first evidence on the gap between micro and macro elasticities, it is however worth noting that our macro estimates are less precisely estimated than our micro ones. Such discrepancy can easily been explained by the limited

[27] In column 7 we show our results when we calculate tax liabilities assuming that all individuals who would be eligible to receive AFDC, TANF or food stamps based on their income actually take-up benefits. Since this leads to larger calculated tax liabilities (and values for the instruments), the estimated marginal effects and elasticities are reduced, but the result that macro participation responses are larger than micro participation responses is actually more pronounced.

policy variations at the state level over time, compared to policy variations across individuals with different number of kids over time. Future research should use other source of policy variations as robustness checks for our macro estimates.

III Simulating the Optimal Tax Schedule

In this Section we show how unemployment and wage responses affect the shape of the optimal tax and transfer schedule. For this purpose we simulate the optimal tax schedule using the sufficient statistics formula for the optimal tax and transfer schedule in the mixed model (see Proposition 3).[28] We calibrate the formula partially relying on the empirical estimates from the previous section and partially relying on parameter estimates from the previous literature. Since there is considerable uncertainty about most of these parameters, these simulations are very stylized and should be viewed as an illustration of the comparative statics of our optimal tax formula, rather than a precise attempt to derive the optimal tax schedule for any particular population.[29]

To simulate the optimal tax schedule, we solve the system of first-order conditions derived in the theoretical section for the tax schedule at different income levels. The system consists of $I+2$ unknowns: the $i = 0 \dots I$ tax levels T_i as well as the Lagrange multiplier λ. These are identified by $I+2$ equations: the I first-order conditions (15), the government budget constraint (1) and finally the normalization $\sum_i h_i g_i = 1$ which holds in the absence of income effects.[30] We base our simulation on the empirical income distribution for single workers in the March 2011 CPS, which we discretize into a grid of bins of \$500 width. Following Saez (2002) we parameterize g_i using the functional form: $g_i = \frac{1}{\lambda (w_i - T_i)^\nu}$, where ν is the parameter describing society's preferences for redistribution. We set $\nu = 0.5$ as our baseline, which corresponds to moderate distributive preferences. Below we will see that this leads to optimal tax schedules similar to the observed schedules and thus may be a good approximation of the preferences underlying U.S. tax policy, but we also show how the optimal tax schedule changes under alternative ν. We also follow Saez (2002) in our specification for the earnings distribution functions $h_i(T)$ and in setting the external revenue requirement of the government E. See Appendix for additional details.

In order to simulate equation (15) we need to specify the following 4 sets of behavioral parameters: The intensive margin mobility elasticity ζ_i, the extensive margin employment elasticity

[28] We thank Emmanuel Saez for sharing his code.

[29] Such an exercise for the U.S. would, for example, have to take into account that policy makers seem to have placed different welfare weights on different groups of single individuals, depending on the number of children. Backing out the implicit welfare weights in the current tax schedule given an optimal tax framework and calibrating how the tax schedule given these welfare weights would change under alternative models would be interesting, but is beyond the scope of this paper.

[30] Alternatively one could also use the FOC for benefit levels derived in the Appendix A.1 which holds in the presence of income effects.

η_i, the employment enhancing effect of taxation μ_i and the matrix ratio of the macro and micro participation responses in equation (16). The intensive margin mobility elasticity ζ_i can be directly mapped into the classical intensive earnings elasticity ε_i using the formula $\zeta_i = \frac{w_i}{w_i - w_{i-1}} \varepsilon_i$.[31] Based on the prior literature, a reasonable range for the intensive margin employment elasticity is $0.2 - 0.3$ (e.g. Chetty et al., 2011a, 2013). We therefore set $\varepsilon_i = 0.25$ for all i.[32] For the extensive margin elasticity, we follow Saez (2002) and set $\eta = 0$ for incomes higher than \$20,000. For incomes lower than \$20,000 we use our empirical estimates from the previous section.

The two key new parameters in our tax formula are μ_i and the matrix ratio of the macro and micro participation responses. Except for two papers (Manning, 1993, Lehmann et al., 2016) that suggests a positive μ_i, we are not aware of empirical estimates of μ_i. We therefore set $\mu_i = 0$ for all i in our baseline and show simulations for a range of possible values for μ_i. For the matrix ratio we start with the case that the matrix ratio is equal to the identity matrix, which corresponds to a situation where micro and macro responses are identical. We then contrast this with a situation where the matrix ratio is still a diagonal matrix, but with diagonal elements different from zero. In particular we use our empirical estimates from the previous section of the macro and micro participation response to calibrate these diagonal elements. This is a practical simplification, as in principle there is no reason to believe that once macro and micro responses are different, the resulting matrix ratio remains diagonal. It is a useful starting point given that this is a relatively straightforward departure from the classic case of equal responses that the prior literature has assumed. In general we believe that estimating macro and micro responses on the intensive margin is a fruitful avenue for future research.

Figure 2 shows the optimal tax and transfer schedule for incomes below \$20,000 and Table 5 provides additional details for these simulations. Panel a) shows post-tax income plotted against pre-tax income and panel b) depicts the corresponding marginal tax rate as a function of pre-tax income. We begin with the case where $\mu_i = 0$ and equal macro and micro participation responses, which thus corresponds exactly to the formula in Saez (2002). The line with circles shows the simulated schedule when we set η_i to our estimate for the micro extensive employment elasticity of 0.6 from Table 2. A key takeaway from our theoretical analysis is that the employment effects should really be measured on the macro level and thus we contrast this scenario to the line with "+" signs to highlight the effect of moving to macro estimates where $\eta_i = 0.42$ consistent with our estimate in Table 2. This moderately increases the benefit for the nonemployed b from \$6,500 to \$7,100. Keeping η_i at the macro estimate level we then use our empirical estimates for the macro and micro participation responses to calibrate the macro/micro matrix ratio to be the diagonal

[31] See Appendix of Saez (2002).

[32] In the on-line appendix we also show simulations for $\varepsilon = 0.5$ for low income workers with qualitatively similar results.

matrix with $0.9 \simeq \frac{-0.028}{-0.031}$ as the element on the diagonal. The resulting tax schedule (shown with squares) now features a substantially higher benefit to the nonemployed of $8,300 and higher marginal tax rates at the bottom. Figure 2b) shows the corresponding marginal tax rates at each income level. This figure highlights the result in Jacquet et al. (2013), that even with large extensive responses, the marginal tax rate at positive incomes is always positive, while there is a discontinuity at 0 where the optimal tax schedule may feature a negative employment tax corresponding to $T_1 + b < 0$. Interestingly, this figure suggests that for these benchmark parameters the optimal tax schedule features an EITC-like structure with a negative employment tax ($T_1 + b < 0$) only as long as we do not take into account the wage and unemployment spillover effects as measured by our (admittedly noisy) gap between macro and micro participation responses.

In Figure 3 we show the same comparisons, but for a wide range of redistributive tastes of the social planner ($\nu = 0.25, 0.5, 1, 4$). The figure highlights that taking into account endogenous wags and unemployment via the macro/micro participation ratio is much more relevant to the shape of the optimal tax schedule for lower values of the redisributive taste parameter, such as $\nu = 0.25$ or $\nu = 0.5$. For higher values of ν our optimal tax formula produces relatively similar results as the Saez (2002) formula. The intuition for this is that our formula essentially rescales the social welfare weights g_i for positive income levels thus putting more weight at the bottom (equation (16)). For high ν the social planner already exhibits essentially Rawlsian preferences and is simply constrained by employment responses to set what is close to the revenue maximizing tax schedule.[33]

Given the considerable uncertainty regarding the macro to micro participation ratio and the value for μ_i, Figure 4 shows comparative statics from our benchmark schedule when varying the macro / micro ratio and μ_i , which we assume to be constant across income groups. Panel (a) shows the tax schedule, setting the ratio to 0.75, 1 and 1.25. For this range we find substantial differences in the benefit b ranging from $9,700 with a low ratio to only $3,200 if the ratio is equal to 1.25. In the latter case, the optimal tax also features a substantial employment credit at the bottom (see Panel (c)). Panel (b) and (d) show the tax schedules for alternative values of μ_i . Higher values of μ_i increase the demogrant and transfers at lower incomes become much larger and shift the break even point (of zero tax liability) further to the right, this in term is financed by substantially higher marginal tax rates at higher income levels. For example, using a value of $\mu = 0.3$ which corresponds to the empirical estimate of Lehmann et al. (2016), the simulations imply a demogrant of $10,400 and an average marginal tax rate of 72 percent. The intuition for this is straightforward: a positive μ_i implies that higher marginal tax rates have an employment enhancing effect by pulling people from zero income into the labor force thus raising new revenue and counteracting the usual intensive margin response that would reduce tax revenue whenever

[33]The on-line appendix provides all simulations from this section for these alternative values for ν as well.

employment tax are positive. The potentially large impact of μ_i on the optimal tax rate and the comparatively sparse empirical literature on this suggest that additional empirical estimates of this parameter would be highly valuable.

Other papers have stressed the possibility that macro employment responses could be significantly lower than micro employment responses, particularly in the context of UI and job search assistance. This has typically been explained with the possibility of job rationing, which may manifest especially during recessions. Our estimates in Table 4, while noisy, are consistent with this view: while both macro and micro responses decline in recessions, the decline is much larger for macro responses, both with respect to employment and participation. The business cycle macro estimates suggest that spillover effects could be larger during economic downturns. Figure 5 simulates how the optimal tax schedule would vary over the business cycle given our estimates from Table 4. We present results from the estimates based on the 6-month change in the unemployment rate, but using the other measures yields qualitatively very similar results. In Figure 5a) and 5c) we show the optimal tax schedule for different business cycle states implied by our optimal tax formula. The transfer at zero income is around \$5,700 during a strong labor market with a large employment credit at the bottom and relatively low marginal tax rates. During weak labor markets the simulation suggests that the transfer at zero should increase to \$10,500 per year with much higher marginal tax rates. In contrast, panels (b) and (d) of Figure 5 show the tax schedule implied by the Saez (2002) formula calibrating η to the macro employment effects estimated over the business cycle.[34] While the decline in macro employment responses during weak labor markets also leads to an increase in transfers at the bottom and a slight increase in employment tax rates, the difference relative to a boom is only around \$500 due to the absence of the spillover channel.

IV Conclusion

This paper revisits the debate about the desirability of the EITC versus the NIT. We have shown that whether the optimal employment tax on the working poor is positive or negative depends on the presence of unemployment and wage responses to taxation. Our sufficient statistics optimal tax formula, combined with our reduced-form empirical estimates, indicate that the optimal policy is pushed more towards an NIT than the standard optimal tax model would suggest, although statistical precision limits strong conclusions about the magnitude of the macro responses.

There are several limitations to our analysis that should be addressed in future work. First, there is clearly a need for better empirical estimates of the macro effects of taxation. Most studies of macro labor supply responses rely on cross-country variation in taxes, which can be substantial. While this variation is clearly desirable for efficiency reasons, across countries, tastes for

[34] Using the micro employment effects yields even less variation in the optimal tax schedule over the cycle.

redistribution and other forms of government spending are probably correlated with taxes and employment and are difficult to fully control for. What is needed is reliable policy variation in taxes across labor markets, similar to variation in UI benefit payments that is exploited in Lalive et al. (2015). Second, it would be very interesting to get better estimates of the wage-moderating effects of tax progressivity. Third, it would be worthwhile to study business cycle effects of taxation more directly by introducing dynamics into the model. The approach we adopted in this paper is entirely steady-state. Finally, it would be useful to develop a model that more fully integrates UI benefits and income taxes, where benefits depend on prior wages, as is currently the policy in most developed economies.

References

Baily, Martin Neil, "Some aspects of optimal unemployment insurance," *Journal of Public Economics*, 1978, *10* (3), 379–402.

Chetty, Raj, "A general formula for the optimal level of social insurance," *Journal of Public Economics*, 2006, *90* (10-11), 1879–1901.

_ , "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

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

_ , Adam Guren, Day Manoli, and Andrea Weber, "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *The American Economic Review*, 2011, *101* (3), 471–475.

_ , _ , _ , and _ , "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," *NBER Macroeconomics Annual*, 2012, *27* (1), 1–56.

_ , John N. Friedman, and Emmanuel Saez, "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, December 2013, *103* (7), 2683–2721.

_ , John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Choné, Philippe and Guy Laroque, "Optimal incentives for labor force participation," *Journal of Public Economics*, 2005, *89* (2-3), 395–425.

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

Currie, Janet and Jonathan Gruber, "Health insurance eligibility and Child Health: lessons from recent expansions of the Medicaid program," *Quarterly Journal of Economics*, 1996, *111* (2), 431–466.

Diamond, Peter, "Income taxation with fixed hours of work," *Journal of Public Economics*, 1980, *13* (1), 101–110.

— , "Optimal Income Taxation: An Example with U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 1998, *88* (1), 83–95.

Eissa, N. and H.W. Hoynes, "Taxes and the labor market participation of married couples: the earned income tax credit," *Journal of Public Economics*, 2004, *88* (9), 1931–1958.

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

— **, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner**, "Evaluation of Four Tax reforms in the United States: Labor supply and welfare effects for single mothers," *Journal of Public Economics*, 2008, *92* (3-4), 795–816.

Gelber, Alexander and Joshua Mitchell, "Taxes and Time Allocation: Evidence from Single Women and Men," *Review of Economic Studies*, 2012, *79*, 863–897.

Golosov, Mikhail, Pricila Maziero, and Guido Menzio, "Taxation and Redistribution of Residual Income Inequality," *Journal of Political Economy*, 2013, *121* (6), 1160 – 1204.

Gruber, Jon and Emmanuel Saez, "The elasticity of taxable income: evidence and implications," *Journal of Public Economics*, 2002, *84* (1), 1–32.

Hansen, Emanuel, "Optimal income taxation with labor supply responses at two margins: When is an Earned Income Tax Credit optimal?," Working Paper, University of Cologne 2018.

Hersoug, Tor, "Union Wage Responses to Tax Changes," *Oxford Economic Papers*, 1984, *36* (1), 37–51.

Hosios, Arthur J, "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 1990, *57* (2), 279–98.

Hoynes, Hilary W. and Ankur J. Patel, "Effective Policy for Reducing Inequality? The Earned Income Tax Credit and the Distribution of Income," NBER Working Papers 21340 2015.

Hummel, Albert Jan, "Unemployment and Tax Design," Working Paper, Erasmus University Rotterdam and Tinbergen Institute 2018.

Hungerbühler, Mathias, Etienne Lehmann, Alexis Parmentier, and Bruno Van der Linden, "Optimal Redistributive Taxation in a Search Equilibrium Model," *Review of Economic Studies*, 2006, *73* (3), 743–767.

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden, "Optimal redistributive taxation with both extensive and intensive responses," 2013, *148* (5), 1770–1805.

_ , _ , and _ , "Optimal income taxation with Kalai wage bargaining and endogenous participation," *Social Choice and Welfare*, 2014, *42* (2), 381–402.

Jäntti, Markus, Jukka Pirttilä, and Håkan Selin, "Estimating labour supply elasticities based on cross-country micro data: A bridge between micro and macro estimates?," *Journal of Public Economics*, 2015, *127* (1), 87 – 99.

Kennan, John and James R. Walker, "Wages, welfare benefits and migration," *Journal of Econometrics*, 2010, *156* (1), 229 – 238. Structural Models of Optimization Behavior in Labor, Aging, and Health.

Kroft, Kory and Matthew J. Notowidigdo, "Should Unemployment Insurance Vary with the Unemployment Rate? Theory and Evidence," 2016, *83* (3), 1092–1124.

Lalive, Rafael, Camille Landais, and Josef Zweimüller, "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 2015, *105* (12), 3564–3596.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez, "A Macroeconomic Approach to Optimal Unemployment Insurance: Applications," *American Economic Journal: Economic Policy*, 2018, *10* (2), 182–216.

_ , _ , and _ , "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory," *American Economic Journal: Economic Policy*, 2018, *10* (2), 152–181.

Lehmann, Etienne, Alexis Parmentier, and Bruno Van der Linden, "Optimal income taxation with endogenous participation and search unemployment," *Journal of Public Economics*, 2011, *95* (11), 1523–1537.

— , **Claudio Lucifora, Simone Moriconi, and Bruno Van der Linden**, “Beyond the Labour Income Tax Wedge: The Unemployment-Reducing Effect of Tax Progressivity,” *International Tax and Public Finance*, 2016, *23* (3), 454–489.

Lockwood, Ben and Alan Manning, “Wage setting and the tax system theory and evidence for the United Kingdom,” *Journal of Public Economics*, 1993, *52* (1), 1–29.

Manning, Alan, “Wage Bargaining and the Phillips Curve: The Identification and Specification of Aggregate Wage Equations,” *Economic Journal*, 1993, *103* (416), 98–118.

Meyer, Bruce D, “Do the poor move to receive higher welfare benefits?,” 2000.

Meyer, Bruce D. and Dan T. Rosenbaum, “Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers,” *The Quarterly Journal of Economics*, 2001, *116* (3), pp. 1063–1114.

Mirrlees, J. A., “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 1971, *38* (2), pp. 175–208.

Moffit, Robert A, “The Effect of Welfare on Marriage and Fertility: What Do We Know and What Do We Need to Know?,” in R. Moffitt, ed., *Welfare, the Family, and Reproductive Behavior*, National Research Concil, National Academy of Sciences Press, 1998.

Mortensen, Dale T. and Christopher A. Pissarides, “New developments in models of search in the labor market,” 1999, *3*, 2567–2627.

Pisauro, Giuseppe, “The effect of taxes on labour in efficiency wage models,” *Journal of Public Economics*, 1991, *46* (3), 329–345.

Pissarides, Christopher A., “Taxes, Subsidies and Equilibrium Unemployment,” *Review of Economic Studies*, 1985, *52* (1), 121–134.

— , “The impact of employment tax cuts on unemployment and wages; The role of unemployment benefits and tax structure,” *European Economic Review*, 1998, *42* (1), 155–183.

Rothschild, Casey and Florian Scheuer, “Redistributive Taxation in the Roy Model,” *The Quarterly Journal of Economics*, 2013, *128* (2), 623–668.

Rothstein, Jesse, “Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence,” *American Economic Journal: Economic Policy*, 2010, *2* (1), pp. 177–208.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin, “Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium,” NBER Working Papers 22646, 2016.

Saez, Emmanuel, "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," *The Quarterly Journal of Economics*, 2002, *117* (3), pp. 1039–1073.

_ , "Direct or Indirect Tax Instruments for Redistribution: short-run versus long-run," *Journal of Public Economics*, 2004, *88*, 503–518.

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

Schmieder, J.F., T. Von Wachter, and S. Bender, "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years," *The Quarterly Journal of Economics*, 2012, *127* (2), 701–752.

Stiglitz, Joseph E., "Self-selection and Pareto efficient taxation," *Journal of Public Economics*, 1982, *17* (2), 213–240.

A Additional Proofs

A.1 Optimal benefit level

The optimal condition with respect to the benefit level is obtained from differentiating (8) with respect to b:

$$\frac{\partial \Lambda}{\partial b} = -1 + \sum_{i=1}^{I} h_i + \sum_{i=1}^{I} (T_i + b) \frac{\partial \mathscr{H}_i}{\partial b} + \frac{u'(b)}{\lambda} \frac{\partial \Omega}{\partial U_0} + \sum_{i=1}^{I} \frac{1}{\lambda} \frac{\partial \Omega}{\partial U_i} \frac{\partial \mathscr{U}_i}{\partial b}$$

Differentiating $\mathscr{U}_i(\mathbf{t}) \equiv \mathscr{P}_i(\mathbf{t})\, u(\mathscr{C}_i(\mathbf{t})) + (1 - \mathscr{P}_i(\mathbf{t}))\, u(b)$ with respect to b gives:

$$\frac{\partial \mathscr{U}_i}{\partial b} = (1 - p_i) u'(b) + p_i\, u'(c_i) \left[\frac{\partial \mathscr{C}_i}{\partial b} + \frac{\partial \mathscr{P}_i}{\partial b} \frac{u(c_i) - u(b)}{p_i\, u'(c_i)} \right]$$

Using $h_0 = 1 - \sum_{i=1}^{I} h_i$ leads to:

$$\frac{\partial \Lambda}{\partial b} = -h_0 + \sum_{i=1}^{I} (T_i + b) \frac{\partial \mathscr{H}_i}{\partial b} + g_0\, h_0 + \sum_{i=1}^{I} g_i\, h_i \left[\frac{\partial \mathscr{C}_i}{\partial b} + \frac{\partial \mathscr{P}_i}{\partial b} \frac{u(c_i) - u(b)}{p_i\, u'(c_i)} \right] \tag{19}$$

where the social marginal welfare weight on the non-employed is:

$$g_0 \stackrel{\text{def}}{\equiv} \frac{u'(b)}{\lambda\ h_0} \left[\frac{\partial \Omega}{\partial U_0} + \sum_{i=1}^{I} (1-p_i) \frac{\partial \Omega}{\partial U_i}\right] \tag{20}$$

We now show how (19) can be simplified if we assume income effect away. In the absence of income effects, a simultaneous change in all tax liabilities and welfare benefit $\Delta T_1 = ... = \Delta T_i = -\Delta b$ induces no changes in wages, conditional employment probabilities not employment levels, so that $\sum_{j=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} = \frac{\partial \mathscr{W}_i}{\partial b}$, $\sum_{j=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_j} = \frac{\partial \mathscr{P}_i}{\partial b}$ and $\sum_{j=1}^{I} \frac{\partial \mathscr{H}_i}{\partial T_j} = \frac{\partial \mathscr{H}_i}{\partial b}$. Using Equations (4) and (5), Equation (10) can be rewritten as:

$$0 = h_j + \sum_{i=1}^{I} (T_i + b) \frac{\partial \mathscr{H}_i}{\partial T_j} + \sum_{i=1}^{I} g_i\ h_i \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i) - u(b)}{p_i\ u'(c_i)}\right] \tag{21}$$

Using that $\frac{\partial \mathscr{C}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial T_i} - 1$ and for $j \neq i$, $\frac{\partial \mathscr{C}_i}{\partial T_j} = \frac{\partial \mathscr{W}_i}{\partial T_j}$, summing (21) for all $j \in \{1, ..., I\}$ and subtracting this sum by $\frac{\partial \Lambda}{\partial b} = 0$ in Equation (19) leads to:

$$\begin{aligned} 0 \quad = \quad & \sum_{i=0}^{I} h_i + \sum_{i=1}^{I} (T_i + b) \left(\sum_{j=1}^{I} \frac{\partial \mathscr{H}_i}{\partial T_j} - \frac{\partial \mathscr{H}_i}{\partial b}\right) - \left(g_0\ h_0 + \sum_{i=1}^{I} g_i\ h_i\right) \\ + \quad & \sum_{i=1}^{I} g_i\ h_i \left(\sum_{j=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} - \frac{\partial \mathscr{W}_i}{\partial b}\right) + \sum_{i=1}^{I} g_i\ h_i \frac{u(c_i) - u(b)}{u'(c_i)} \left(\sum_{j=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_j} - \frac{\partial \mathscr{P}_i}{\partial b}\right) \end{aligned} \tag{22}$$

In the absence of income effects, (22) can be simplified to:

$$g_0\ h_0 + \sum_{i=1}^{I} g_i\ h_i = \sum_{i=0}^{I} h_i = 1$$

A.2 Derivation of Equation (15)

Using Equation (16), Equation (11) can be rewritten as:

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathscr{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) - (\hat{g}_1\ h_1, ... \hat{g}_I\ h_I)^T$$

That is, for each $j = 1, .., I$, as:

$$0 = (1 - \hat{g}_j) h_j + \sum_{i=1}^{I} (T_i + b) \frac{\partial \mathscr{H}_i}{\partial T_j}$$

Using (14a), we get:

$$\begin{aligned}(1-\hat{g}_j)h_j &= -(T_{j-1}+b)\frac{\partial \mathscr{H}_{j-1}}{\partial T_j} - (T_j+b)\frac{\partial \mathscr{H}_j}{\partial T_j} - (T_{j+1}+b)\frac{\partial \mathscr{H}_{j+1}}{\partial T_j} \\ &= -(T_{j-1}+b)\frac{\partial \widetilde{\mathscr{H}_{j-1}}}{\partial (T_j - T_{j-1})} - (T_j+b)\left[\frac{\partial \widetilde{\mathscr{H}_j}}{\partial (T_j - T_{j-1})} + \frac{\partial \widetilde{\mathscr{H}_j}}{\partial T_j} - \frac{\partial \widetilde{\mathscr{H}_j}}{\partial T_{j+1} - T_j}\right] + (T_{j+1}+b)\frac{\partial \widetilde{\mathscr{H}_{j+1}}}{\partial T_{j+1} - T_j}\end{aligned}$$

And so, using (14d)

$$\begin{aligned}\left(1-\hat{g}_j - \frac{T_j+b}{c_j-b}\eta_j\right)h_j &= -\left[(T_{j-1}+b)\frac{\partial \widetilde{\mathscr{H}_{j-1}}}{\partial (T_j - T_{j-1})} + (T_j+b)\frac{\partial \widetilde{\mathscr{H}_j}}{\partial (T_j - T_{j-1})}\right] \\ &+ \left[(T_j+b)\frac{\partial \widetilde{\mathscr{H}_j}}{\partial (T_{j+1} - T_j)} + (T_{j+1}+b)\frac{\partial \widetilde{\mathscr{H}_{j+1}}}{\partial (T_{j+1} - T_j)}\right]\end{aligned}$$

Summing the latter equality for $j = i, ..., I$ leads to:

$$-\left[(T_{i-1}+b)\frac{\partial \widetilde{\mathscr{H}_{i-1}}}{\partial (T_i - T_{i-1})} + (T_i+b)\frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})}\right] = \sum_{i=j}^{I}\left(1-\hat{g}_j - \frac{T_j+b}{c_j-b}\eta_j\right)h_j$$

Combining the latter equation with

$$\begin{aligned}-\left[(T_{i-1}+b)\frac{\partial \widetilde{\mathscr{H}_{i-1}}}{\partial (T_i - T_{i-1})} + (T_i+b)\frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})}\right] &= -(T_i - T_{i-1})\frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})} \\ &- (T_{i-1}+b)\left[\frac{\partial \widetilde{\mathscr{H}_{i-1}}}{\partial (T_i - T_{i-1})} + \frac{\partial \widetilde{\mathscr{H}_i}}{\partial (T_i - T_{i-1})}\right]\end{aligned}$$

and with Equations (14b) and (14c) leads to (15).

Table 1: Variable Means for Single Individuals

	(1) Estimation Sample	(2) High School Dropout	(3) High School Graduate	(4) Some College
Panel A: Demographics				
Age	32.9	32.3	32.7	33.5
No Children Percent	79.4	77.1	80.2	79.6
1 Child Percent	10.9	10.3	10.8	11.5
2 Children Percent	6.2	6.8	6.0	6.2
3+ Children Percent	3.4	5.8	3.0	2.7
Mean Years of Education	11.9	9.2	12	13.3
Percent Black	17.5	20.6	17.7	15.5
Percent Hispanic	14.2	29.2	11.5	9.5
Panel B: Labor Force Status				
Labor Force Participation Rate (k_i)	80.8	65.2	82.7	87.0
Employment Rate (h_i)	72.9	54.1	74.5	81.1
Unemployment Rate $(1-p_i)$	10.3	17.2	10.0	6.8
Panel C: Income, Taxes and Transfers (Real 2010 Dollars)				
Imputed Pre-tax Wage Earnings	19268	11878	19197	23456
Real Total post-tax and transfer income with takeup: ols	15618	11067	15513	18278
Net Taxes: No Children	4674	2349	4638	5968
Net Taxes: 2 Children	-76	-1633	-400	1292
AFDC/TANF and Food Stamps: No Children	499	1086	440	261
AFDC/TANF and Food Stamps: 2 Children	3476	6805	3245	1726
Net Tax and Transfers (T_i): No Children	4175	1264	4197	5707
Net Tax and Transfers (T_i): 2 Children	-3552	-8438	-3645	-435
Net Tax and Transfers (b): Zero Income, No Children	-2071	-2055	-2073	-2078
Net Tax and Transfers (b): Zero Income, 2 Children	-11646	-11693	-11640	-11627
AFDC/TANF Recipiency Rate for Mothers: Pre-1996	25	43	21	15
AFDC/TANF Recipiency Rate for Mothers: Post-1996	9	17	8	5
Food Stamp Recipiency Rate: Pre-1996	12	25	11	6
Food Stamp Recipiency Rate: Post-1996	14	25	14	10
Number of observations	1817083	350817	832919	633347

Notes: The sample is restricted to single men and women aged 18-55. All dollar figures are in real 2010 dollars. Data used in each column are restricted to individuals with the education level in the column header.
Imputed earnings result from a linear regression of demographics on wages conditional on employment.
Net Taxes is federal, state and fica (sum of employer and employee) tax liabilities net of tax credits, including EITC.
AFDC/TANF and Food Stamps assume 100 percent recipiency among those eligible based on income.
Net Taxes and Transfers is the net of federal, state and fica (sum of employer and employee) tax liabilities and credits, AFDC or TANF payments and food stamp benefits.

Table 2: Micro and Macro Responses to Changes in Taxes and Benefits
Instrumental Variable Regressions

LHS Variable	(1) Participation Rate: $\hat{\mathcal{K}}_i$	(2) Employment Rate: $\hat{\mathcal{H}}_i$
Micro Response		
Taxes Plus Benefits	-0.031	-0.029
	[0.002]***	[0.002]***
Num. Obs	1816065	1816065
Mean of Dep. Var.	0.81	0.73
Income Gain from Employment (2010USD)	15014	15014
Tax Elasticity	-0.57	-0.60
Macro Response: $\frac{\partial}{\partial T_i}$		
Avg Taxes Plus Benefits within Labor Market	-0.028	-0.022
	[0.014]**	[0.021]
Num. Obs	8568	8568
Mean of Dep. Var.	0.79	0.70
Income Gain from Employment (2010USD)	13664	13664
Tax Elasticity	-0.48	-0.42

Notes: (* P<.1, ** P<.05, *** P<.01) Standard errors clustered on state level. The sample is restricted to single women aged 18-55. The data include March CPS for 1984-2011 and Outgoing Rotations Groups for 1994-2010. The first column uses labor force participation as the outcome variable, the second column uses employment status. Taxes Plus Benefit is the net of federal (including EITC), state and fica (sum of employer and employee) taxes plus the benefits an individual would be eligible for at no earnings, adjusted for national recipiency rates. The Micro Response regressions use individual level data and include controls for age, age-squared, race, ethnicity and fixed effects for number of children and State x Year x Month fixed effects, all interacted with education. The Macro Response regressions use data that are collapsed to the state-year cell, each cell receives equal weight in the regression. Regressions include controls (all interacted with education) for percent black, percent hispanic, average age, age-squared, number of children and fixed effects for state and year and CPS region time trends.

Table 3: Alternative Estimates of Participation and Employment Responses

	(1) Region Time Trend	(2) Div X Year FE	(3) Reg X Year FE	(4) No Pre-Trends	(5) No State Taxes	(6) State- Unemp.	(7) Full Take-up
Micro Participation Response							
Taxes Plus Benefits	-0.031	-0.031	-0.031	-0.031	-0.034	-0.032	-0.018
	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.001]***
Num. Obs	1816065	1816065	1816065	1816065	1816065	1816065	1816065
Mean of Dep. Var.	0.81	0.81	0.81	0.81	0.81	0.81	0.81
Income Gain from Empl.	15014	15014	15014	15014	14501	15014	15475
Tax Elasticity	-0.57	-0.57	-0.57	-0.57	-0.61	-0.57	-0.36
Macro Participation Response							
Avg Taxes Plus Benefits	-0.028	-0.030	-0.031	-0.034	-0.028	-0.031	-0.009
within Labor Market	[0.014]**	[0.018]	[0.015]*	[0.016]**	[0.022]	[0.013]**	[0.006]
Num. Obs	8568	8568	8568	8568	8568	8568	8568
Mean of Dep. Var.	0.79	0.79	0.79	0.79	0.79	0.79	0.79
Income Gain from Empl.	13664	13664	13664	13664	12695	13664	13914
Tax Elasticity	-0.48	-0.51	-0.52	-0.53	-0.48	-0.52	-0.13
Macro Employment Response							
Avg Taxes Plus Benefits	-0.022	-0.023	-0.021	-0.032	-0.034	-0.030	-0010
within Labor Market	[0.021]	[0.022]	[0.020]	[0.019]	[0.022]	[0.015]**	[0.009]
Num. Obs	8568	8568	8568	8568	8568	8568	8568
Mean of Dep. Var.	0.70	0.70	0.70	0.70	0.70	0.70	0.70
Income Gain from Empl.	13664	13664	13664	13664	13664	12479	13914
Tax Elasticity	-0.42	-0.43	-0.41	-0.52	-0.53	-0.49	-0.20

Notes: (* P<.1, ** P<.05, *** P<.01) Standard errors clustered on state level. The sample is restricted to single men and women aged 18-55. The data includes March CPS for 1984-2011 and Outgoing Rotations Groups for 1994-2010. Our baseline specification from Table 3 is contained in column (1). Column (2) replaces region-specific linear time trends with division-by-year fixed effects. Column (3) replaces region-specific linear time trends with region-by-year fixed effects. Column (4) drops region-specific linear time trends. Column (5) is our baseline specification but drops state taxes, including state EITC supplements, from both the OLS and IV tax liabilities. Taxes Plus Benefit is the net of federal (including EITC), state and fica (sum of employer and employee) taxes plus the benefits an individual would be eligible for at no earnings, adjusted for national recipiency rates. Column (6) controls for the state unemployment rate interacted with education. Column (7) is our baseline specification but assumes 100 percent take-up rates for AFDC/TANF and Food Stamps for the computation of the imputed tax liability and the simulated instrument.

Table 4: Participation and Employment Responses: Heterogeneous Labor Market Conditions

	(1)	(2)	(3)	(4)
	Regression Coef.		**Extrapolated Marg. Effects**	
	Marginal Effect of Tax Liability	Interaction of Tax Liab. with Labor Market Meas.	Weak Labor Market	Strong Labor Market
Panel A: Micro Participation				
6-mo change in unemp	-0.031	0.0011	-0.030	-0.033
	[0.002]***	[0.0004]***		
State unemp. rate	-0.032	0.0012	-0.027	-0.036
	[0.002]***	[0.0003]***		
Unemp above 9 pct	-0.032	0.0054	-0.026	-0.032
	[0.002]***	[0.0013]***		
Panel B: Macro Participation				
6-mo change in unemp	-0.028	0.0046	-0.022	-0.035
	[0.015]*	[0.0028]*		
State unemp. rate	-0.031	0.0011	-0.026	-0.036
	[0.017]*	[0.0012]		
Unemp above 9 pct	-0.030	0.0090	-0.021	-0.030
	[0.016]*	[0.0052]*		
Panel C: Micro Employment				
6-mo change in unemp	-0.029	0.0007	-0.028	-0.030
	[0.002]***	[0.0005]		
State unemp. rate	-0.029	0.0015	-0.024	-0.035
	[0.002]***	[0.0003]***		
Unemp above 9 pct	-0.029	0.0074	-0.022	-0.029
	[0.002]***	[0.0018]***		
Panel D: Macro Employment				
6-mo change in unemp	-0.022	0.0030	-0.018	-0.025
	[0.018]	[0.0031]		
State unemp. rate	-0.029	0.0018	-0.021	-0.036
	[0.018]*	[0.0013]		
Unemp above 9 pct	-0.028	0.0112	-0.018	-0.029
	[0.016]*	[0.0060]*		

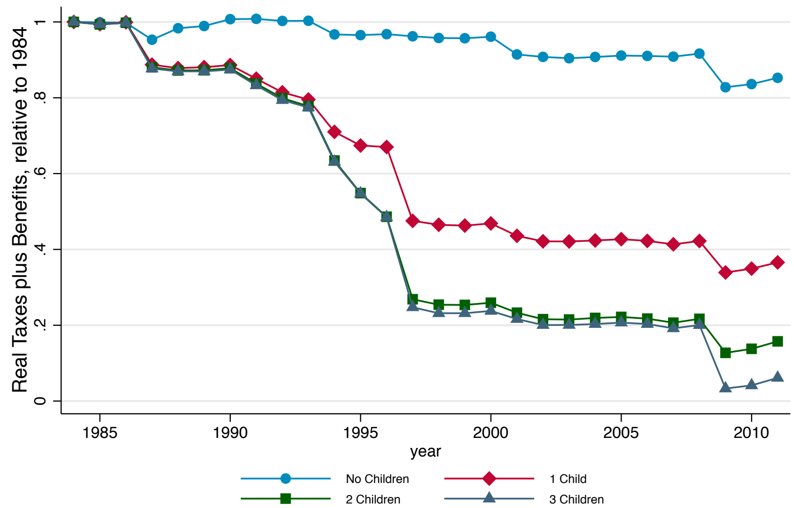
Notes: (* P<.1, ** P<.05, *** P<.01) Standard errors clustered on state level. The Micro Response regressions use individual level data and include controls for age, age-squared, race, ethnicity and fixed effects for number of children and State x Year x Month. The Macro Response regressions use data that are collapsed to the state-year cell observations, each cell receives equal weight in the regression. Regressions include controls for percent black, percent hispanic, average age, age-squared, number of children and fixed effects for state and year and CPS Region time trends. Weak and strong labor markets marginal effects assume the market indicator is two standard deviations above or below the mean for the continuous variables.

Table 5: Simulating the Optimal Tax and Transfer Schedule - $\nu = 0.50$

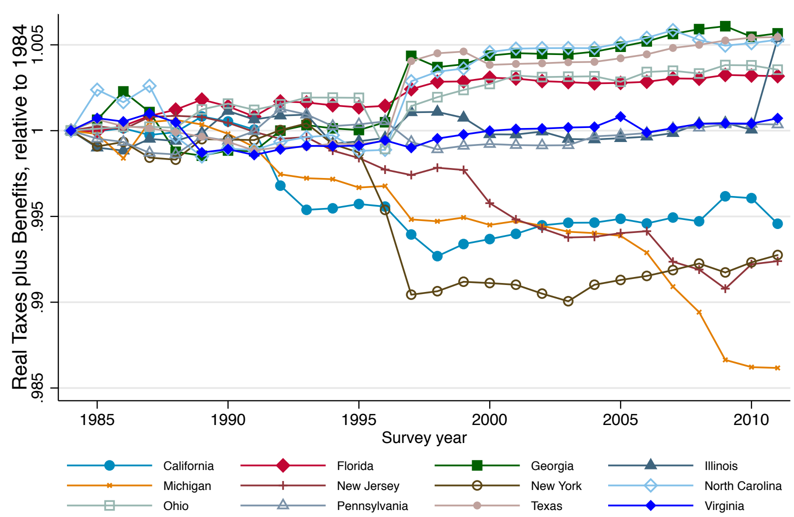
	η	μ	$\frac{\frac{d\mathcal{K}}{dT}}{\frac{d\mathcal{K}}{dT}\vert^{Micro}}$	Demogrant b	Emp. Tax $T_1 + b$	ATR (6k-0)	Avg. MTR	Break-even $T(w) = 0$
Panel A: Alternative η - Macro/Micro ratio = 1, $\mu = 0$								
	0.01	0.00	1.00	9300	500	84.89	0.56	11900
Macro Est.	0.42	0.00	1.00	7100	-200	40.21	0.51	13800
Micro Est.	0.60	0.00	1.00	6500	-300	29.11	0.50	14400
	1.00	0.00	1.00	5600	-400	14.10	0.46	15500
Panel B: Comparing different Macro-Micro Participation Ratios								
	0.42	0.00	0.75	9700	300	69.01	0.62	13500
Benchmark	0.42	0.00	0.90	8300	100	54.75	0.56	13600
	0.42	0.00	1.00	7100	-200	40.21	0.51	13800
	0.42	0.00	1.25	3200	-3200	-17.71	0.44	14500
Panel C: Comparing different values for μ								
	0.42	0.00	1.00	7100	-200	40.21	0.51	13800
	0.42	0.10	1.00	8000	-300	32.98	0.57	15700
	0.42	0.20	1.00	9100	-300	24.57	0.63	17900
	0.42	0.30	1.00	10400	-300	15.52	0.72	20500
Panel D: Optimal Tax Schedule over Business Cycle - with Wage and Unemployment Responses								
Recession	0.34	0.00	0.73	10100	300	74.49	0.63	13400
Normal	0.42	0.00	0.90	8300	100	54.75	0.56	13600
Boom	0.48	0.00	1.06	6000	-700	24.30	0.48	14100
Panel E: Optimal Tax Schedule over Business Cycle - without Wage and Unemployment Responses								
Recession	0.34	0.00	1.00	7400	-100	46.12	0.52	13500
Normal	0.42	0.00	1.00	7100	-200	40.21	0.51	13800
Boom	0.48	0.00	1.00	6900	-300	36.30	0.51	14000

Notes: The table shows simulations of the optimal tax schedule based on equation 15 in the text. All simulations are based on the single worker earnings distribution in the March 2011 CPS and assume the parameter values $\nu = 0.50$, $\varepsilon = 0.25$ and - for wages above \$20,000 - $\eta = 0$. In Pandel D, both η and the Macro/micro ratio vary between boom and recession, while in Panel E the Macro/micro ratio is held constant at 1.

Figure 1: The Variation in Taxes plus Benefits



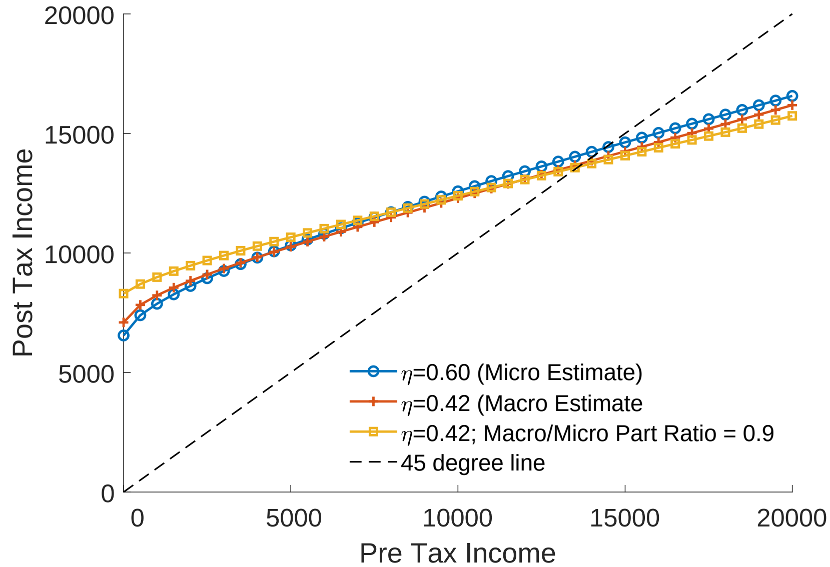
(a) Micro Variation in Taxes plus Benefits



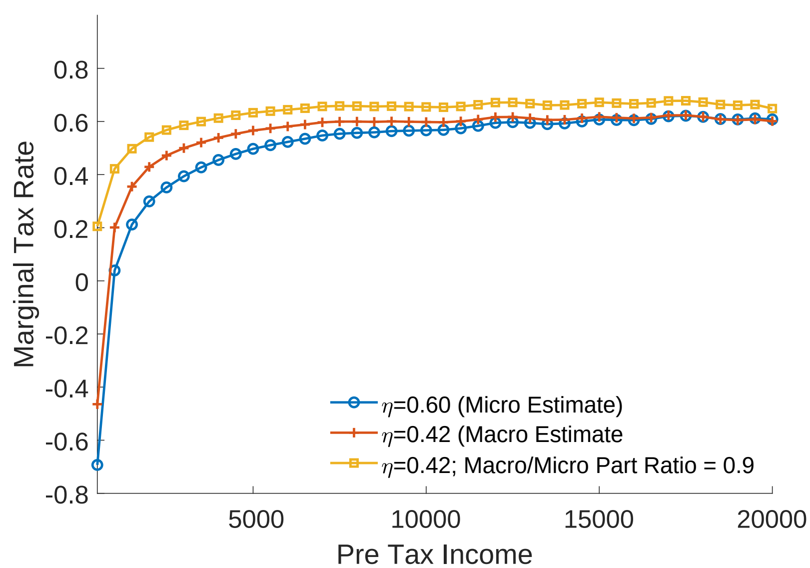
(b) Macro Variation in Taxes plus Benefits

Notes: The top figure shows the variation in taxes plus benefits for high school dropouts by number of children normalized such that 1984 equals one. Taxes plus benefits is the net of federal (including EITC), state and fica (sum of employer and employee) taxes plus the benefits an individual would be eligible for at no earnings, adjusted for national recipiency rates.
The bottom figure shows residuals from a regression of year fixed effects on the state level average taxes plus benefits with state means added back to the residual, then normalized such that 1984 equals one. Taxes plus benefits is the net of federal (including EITC), state and fica (sum of employer and employee) taxes plus the benefits an individual would be eligible for at no earnings, adjusted for national recipiency rates.

Figure 2: Optimal Tax and Transfer Schedule - Baseline



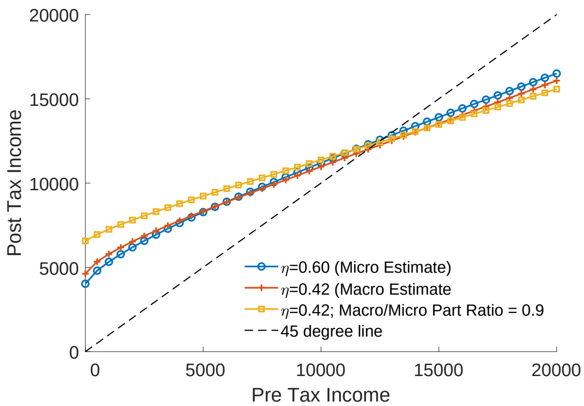
(a) Post vs. Pre-tax income, $\nu = 0.5$



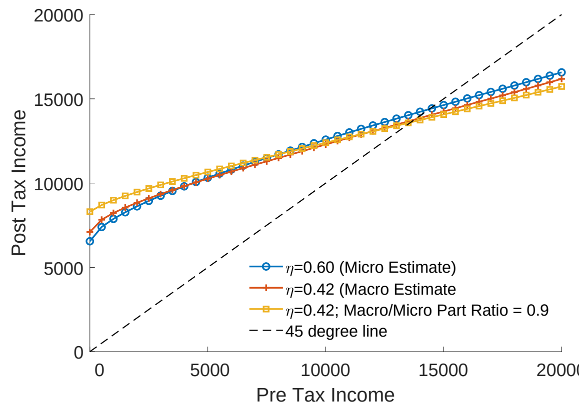
(b) Marginal tax rates, $\nu = 0.5$

Notes: Simulations of the optimal tax and transfer schedule under alternate assumptions on employment and participation responses using eqn (15) in the paper. All simulations assume that the intensive margin elasticity $\varepsilon = 0.25$ at all income levels and that the employment enhancing effect of taxation $\mu = 0$ at all income levels. The first two lines (blue with circle and red with plus signs) assume that the Macro/Micro particiipation response ratio is equal to 1 and sets the extensive margin employment elasticity to 0.6 and 0.42 respectively. The yellow line with stars sets $\eta = 0.42$ and the Macro/Micro participation response ratio equal to the empirical estimate of 0.9. In Panel b) the first dot of each line is the employment tax rate of moving from 0 income to the first income bin.

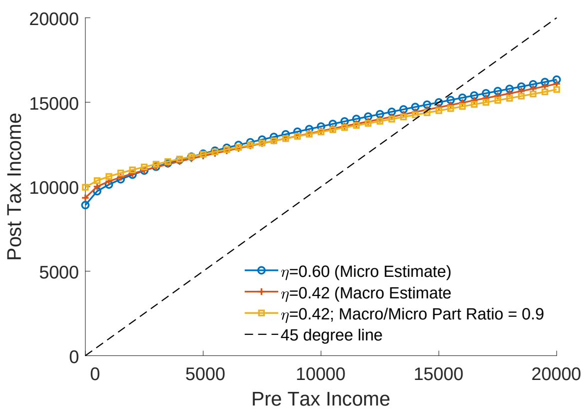
Figure 3: Optimal Tax and Transfer Schedule under alternative Redistributive Preferences



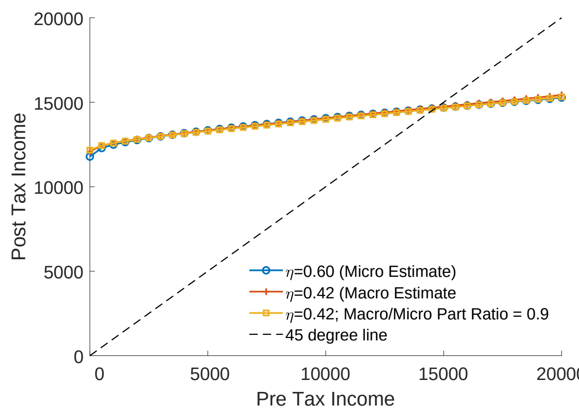
(a) Low redistributive taste parameter: $\nu = 0.25$



(b) Medium-low redistributive taste parameter: $\nu = 0.5$



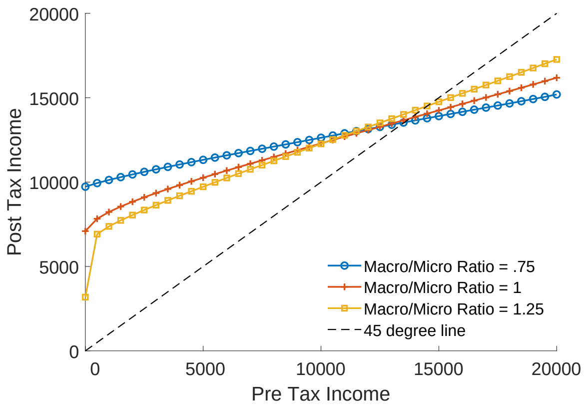
(c) Medium-high redistributive taste parameter: $\nu = 1$



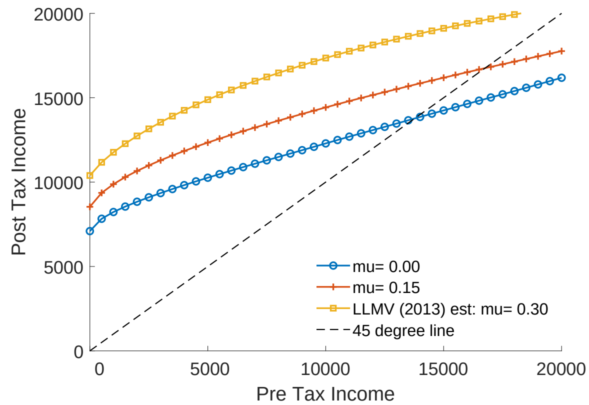
(d) High redistributive taste parameter: $\nu = 4$

Notes: Simulations of the optimal tax and transfer schedule using eqn (15) in the paper. The simulations correspond to Figure 2, but each panel uses different redistributive taste parameters ν. All simulations assume: $\varepsilon = 0.25$, $\eta_{low} = 0.42$, $\eta_{high} = 0$ and $\mu = 0$.

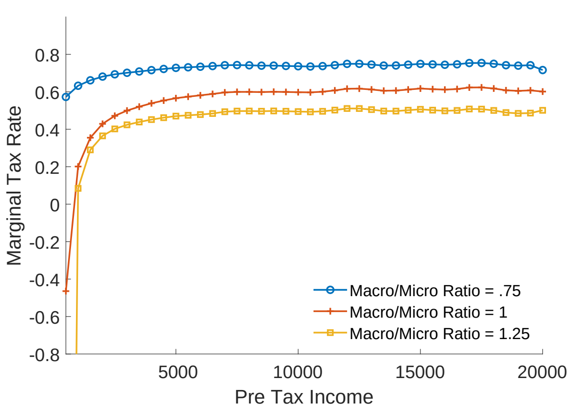
Figure 4: The Effect of the Macro-Micro Participation Ratio and μ on the Optimal Tax and Transfer Schedule



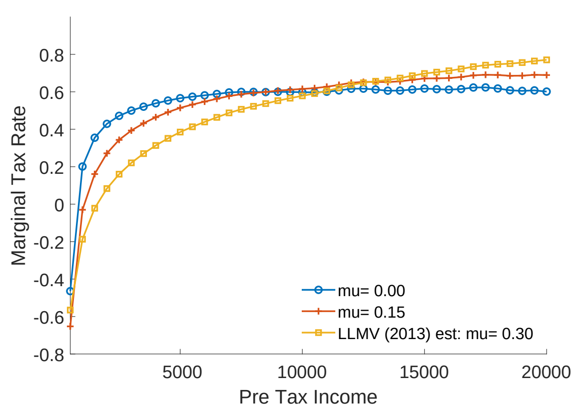
(a) Post vs. Pre-tax income, Alternative Macro/Micro ratio



(b) Post vs. Pre-tax income, Alternative μ



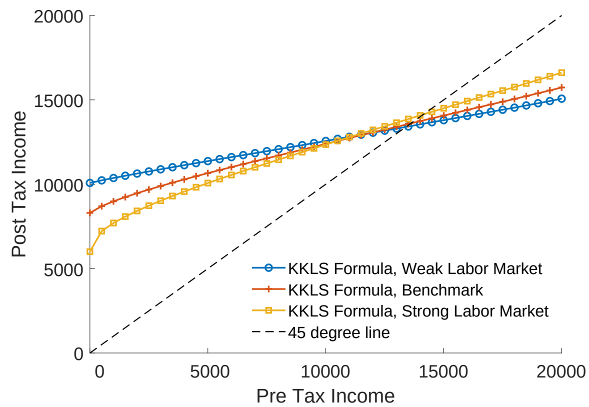
(c) Marginal tax rates, Alternative Macro/Micro ratio



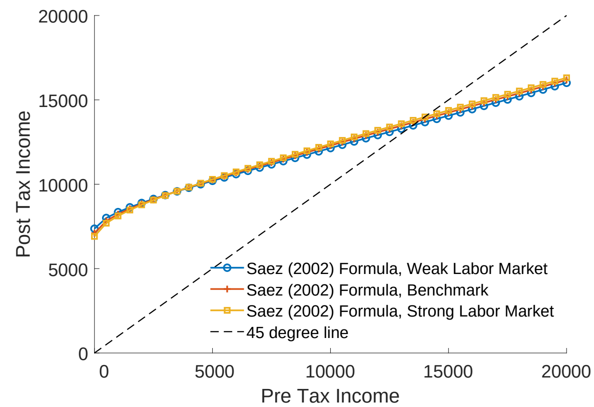
(d) Marginal tax rates, Alternative μ

Simulations of the optimal tax and transfer schedule using eqn (15) in the paper. Panels a) and c) contrast the optimal tax and transfer schedule under different values for the Macro/Micro participation response ratio while setting $\mu = 0$. Panels b) and d) contrast the optimal schedule for different values of μ while setting the Macro/Micro participation response ratio to 1. All simulations assume: $\varepsilon = 0.25$, $\eta_{low} = 0.42$, $\eta_{high} = 0$ and the redistributive taste parameter $\nu = 0.5$. In Panel c) and d) the first dot of each line is the employment tax rate of moving from 0 income to the first income bin.

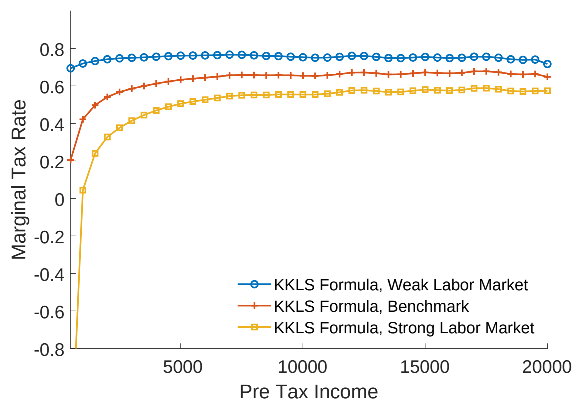
Figure 5: Optimal Tax and Transfer Schedule in Weak vs. Strong Labor Markets



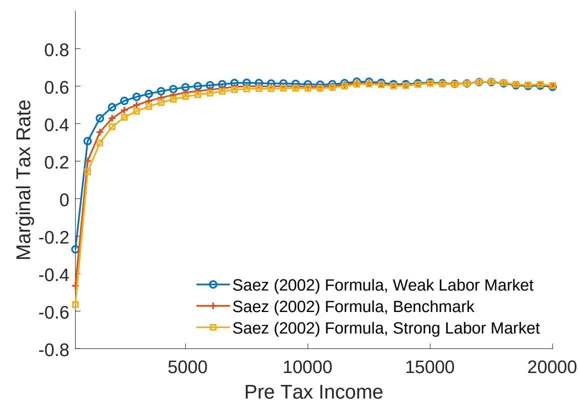
(a) KKLS formula: Post vs. Pre-tax income



(b) Saez (2002) formula: Post vs. Pre-tax income



(c) KKLS formula: Marginal tax rates



(d) Saez (2002) formula: Marginal tax rates

Notes: Simulations of the optimal tax and transfer schedule using eqn (15) in the paper. Panels a) and c) contrast the optimal tax and transfer schedule during weak and strong labor markets using the empirical estimates in the paper. Panels a) and c) use the optimal tax formula in the paper where both the employment elasticity η_{low} varies between boom and recessions as well as the Macro/Micro participation response. Panels b) and d) set the Macro/Micro participation response ratio to 1 and let's only the employment elasticity vary over the cycle and thus corresponds to the Saez (2002) formula . The solid line shows the tax schedule using the weak labor market estimates from Table 4 based on the 6 month change in the unemployment rate. The line with plus signs shows the tax schedule for the corresponding strong labor market estimates from Table 4. All simulations assume $\varepsilon = 0.25$, $\eta_{high} = 0$, $\mu = 0$ and the redistributive taste parameter $\nu = 0.5$. In Panel c) and d) the first dot of each line is the employment tax rate of moving from 0 income to the first income bin.