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

Kory KROFT†
University of Toronto, NBER

Kavan KUCKO‡
Boston University

Etienne LEHMANN§
CRED, University Panthéon-Assas Paris II

Johannes SCHMIEDER¶
Boston University, NBER and IZA

July 19, 2015

Abstract

This paper reassesses whether the optimal income tax program features an Earned Income Tax Credit (EITC) or a Negative Income Tax (NIT) at the bottom of the income distribution, in the presence of unemployment and wage responses to taxation. The paper makes two key contributions. First, it derives a sufficient statistics optimal tax formula in a general model that incorporates unemployment and endogenous wages. This formula nests a broad variety of structures of the labor market, such as competitive models with fixed or flexible wages and models with matching frictions. Our results show that the sufficient statistics to be estimated are: the macro employment response with respect to taxation and the micro and macro participation responses with respect to taxation. We show that an EITC-like policy is optimal provided that the welfare weight on the working poor is larger than the ratio of the micro participation elasticity to the macro participation elasticity. The second contribution is to estimate the sufficient statistics that are inputs to the optimal tax formula using a standard quasi-experimental research design. We estimate these reduced-form parameters using policy variation in tax liabilities stemming from the U.S. tax and transfer system for over 20 years. Using our empirical estimates, we implement our sufficient statistics formula and show that the optimal tax at the bottom more closely resembles an NIT relative to the case where unemployment and wage responses are not taken into account.

*We would like to thank Felix Bierbrauer, Pierre Cahuc, David Card, Raj Chetty, Sebastian Findeisen, Peter Funk, Robert Gary-Bobo, Emmanuel Hansen, Henrik Kleven, Claus Thustrup Kreiner, Patrick Kline, Kevin Lang, Thomas Le Barbanchon, Isabelle Mejean, Jean-Baptiste Michau, Austin Nichols, Matthew Notowidigdo, Claudia Olivetti, Daniele Paserman, Julien Prat, Jesse Rothstein, Dominik Sacks, Emmanuel Saez, Stefanie Stantcheva, Aleh Tsyvinski, Owen Zidar as well as seminar participants at Boston University, McGill University, University of Connecticut, University of Toronto, CRED-University Paris II, CREST, TEPP Aussois winter school, University of Cologne, Queens University, SOLE Meeting 2013, the NTA Meeting 2014, Berkeley, Hamilton, Wilfried Laurier, CEPR, SED 2015 Warsaw 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, 75 005 Paris, France. Etienne Lehmann is also research fellow at IDEP, IZA and CESifo and is junior member of IUF.

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

I Introduction

Recent decades have witnessed a large shift in the U.S. tax and transfer system away from welfare towards in-work benefits. In particular, for single mothers, work incentives increased dramatically: welfare benefits were cut and time limits introduced, the Earned Income Tax Credit (EITC) was expanded and changes in Medicaid, job training programs and child care provision encouraged work. The shift away from programs featuring a Negative Income Tax (NIT) structure (lump-sum transfers to the non-employed with positive employment taxes) towards EITC-like programs (negative employment taxes at the bottom) is prevalent in other countries including Canada, France, South Korea and the U.K.

The literature evaluating these policy reforms largely views them as successful. For single mothers, the reforms sharply reduced welfare caseloads and increased labor force participation (Eissa and Liebman, 1996, Meyer and Rosenbaum, 2001, Eissa and Hoynes, 2006, Gelber and Mitchell, 2012) and consumption levels (Meyer and Sullivan, 2004, 2008). Within an optimal income taxation framework, the various tax policy changes substantially improved welfare (Eissa et al., 2008). This is consistent with (Saez, 2002) who shows that the optimal income tax features an EITC-like structure at the bottom of the income distribution when labor supply responses are primarily concentrated along the extensive margin relative to the intensive margin and the social welfare weight on the working poor exceeds one.

Two important assumptions in (Eissa et al., 2008) and (Saez, 2002) are that *all* job-seekers find work and wages are fixed with respect to the tax system. The first assumption may be appropriate during the 1990s when the unemployment rate was falling and was very low, by historical standards but may be less realistic in more recent periods where unemployment rates exceeded 10 percent. Furthermore, even in a full employment economy, the assumption of fixed wages may be implausible (Rothstein, 2010). The goal of this paper is to reassess whether the optimal income tax features an EITC-like structure at the bottom, in the presence of unemployment and wage responses to taxation.

The paper makes two contributions. First, it derives a sufficient statistics optimal tax formula in a general model that incorporates unemployment and wage responses. This addresses (Mirrlees, 1999) who writes that "a desire is to have a model in which unemployment can arise and persist for reasons other than a preference for leisure". In the model, individuals can be out of work by choice ("non-participants") or by failing in their search to find a job ("unemployed"). This contrasts with the standard model where all active individuals are effectively working. Rather than specifying the structure of the labor market, we pursue a sufficient statistics approach (Chetty, 2009) by allowing wages and the "conditional employment probability" - the rate at which workers in the labor force are employed (i.e. one minus the unemployment rate) - to depend in a reduced-form way on taxes. Competitive models with fixed and flexible wages (Diamond (1980), Saez (2002, 2004), Choné and Laroque (2005), Choné and Laroque (2011), Rothstein (2010) and Lee and Saez (2012)) and models with matching frictions (Hungerbühler et al. (2006) and Landais et al. (2015)) are special cases of our sufficient statistics formula. Thus, a primary advantage of our approach is its generality with respect to the underlying mechanisms. Also, our formulas are exact and do not rely on any approximations. Since our optimal tax formulas are stated in terms of estimable policy parameters, they can be implemented empirically. The disadvantage of our approach however is that analytical results about the precise shape of the optimal tax schedule are harder to obtain and it is more difficult to conduct counterfactual policy experiments.

Our theoretical results show that, for each labor market, the sufficient statistics to be estimated are: *i*) the microeconomic participation response with respect to taxation, *ii*) the macroeconomic participation response with respect to taxation and *iii*) the macroeconomic employment response with respect to taxation.[1] Unlike the micro elasticity, the macro elasticities allow wages and the conditional employment probabilities in each labor market to respond to taxes. We show that an EITC-like policy is optimal provided that the welfare weight on the working poor is larger than the ratio of the micro participation elasticity to the macro participation elasticity. When the micro and macro effects are equal, this collapses to the condition in (Saez, 2002). Thus, if the macro effect is less than the micro effect, as our empirical evidence suggests, the optimal policy is pushed more towards an NIT, relative to the benchmark case.

The intuition for why our optimal tax formula depends on macro employment responses and on the ratio (in matrix terms) of the macro and micro participation responses is the following. In the absence of unemployment and wage responses, behavioral responses to taxation only matter through their effects on the government's budget because they have no first-order effect on an individual's objective by the envelope theorem (Saez, 2001, 2002). However, the latter argument does not apply to wage and unemployment responses because these responses are not directly chosen by individuals but rather are mediated at the market level.[2] Since the social welfare function is assumed to depend only on expected utilities, market spillovers due to wage and unemployment responses matter only insofar as macro responses of expected utility to taxes differ from micro responses. Moreover, since participation decisions depend only on expected utility as well, these market spillovers are entirely captured by the ratio of macro over micro participation responses. This is related to results in Kroft (2008) and Landais et al. (2015) who show that to evaluate optimal unemployment insurance (UI), it is important to estimate the ratio of the micro and macro take-up

[1]For ease of exposition, we hereafter refer to microeconomic as ”micro” and macroeconomic as ”macro”.

[2]For example, higher taxes in one occupation may change equilibrium wages, and therefore labor demand of firms and the conditional employment probabilities that workers face. Such responses may also appear in occupations other than the one where the tax has changed. Moreover, the tax change may reduce the number of job seekers, thereby triggering search externalities

and duration elasticities in the presence of spillover effects, respectively.

The second contribution of this paper is to estimate the sufficient statistics that are inputs to the optimal tax formula using a standard quasi-experimental research design. We contribute to the empirical tax literature in three ways. First, many studies in the tax literature do not clarify whether labor supply responses correspond to micro or macro elasticities. An important exception is (Rothstein, 2010) who considers labor demand responses to the EITC in the U.S. Like Rothstein (2010), our empirical work emphasizes this important distinction. Additionally, we estimate micro and macro effects for multiple skill groups, which is necessary to implement our optimal tax formula, and we use a single methodology and the same sample.[3] This avoids the concern that differences in micro and macro estimates are confounded by differences in methodologies and/or different samples. Second, our results clarify the importance of distinguishing between the effects of taxes on labor force participation and employment. Some studies use the labor force participation rate as the dependent variable (Gelber and Mitchell, 2012) while others use the employment rate (Meyer and Rosenbaum, 2001). Our optimal tax formula indicates that it is important to estimate *both* participation and employment elasticities. Third, this study adds to the large literature evaluating the impact of the EITC expansions in the 1980s and 1990s by expanding the analysis horizon until the most recent years.[4]

Following most of the empirical tax literature, we focus on single women. The primary advantage is that this group is most likely to be at the margin of participating in the labor market and is thereby most affected by tax and transfer policies at the bottom of the income distribution, in particular the EITC.[5] We use data for 1984-2011 from the Current Population Survey (CPS) to measure labor force participation, employment and the number of children in the household. To measure welfare benefits, we use the Urban Institute welfare calculator which contains information on Aid to Families with Dependent Children (AFDC), Temporary Assistance for Needy Families (TANF) and Supplemental Nutrition Assistance Program (SNAP). To measure taxes, we use NBER TAXSIM which contains state and federal income and payroll taxes (including the EITC).

We adopt a "cell-based" approach and define labor markets on the basis of education (high school dropouts, high school graduates, some college but no degree, and college graduates), state

[3] A recent study by Jäntti et al. (2015) estimates micro and macro labor supply elasticities using cross-country data from the Luxembourg Income Study (LIS) along with a single estimator.

[4] One of the earliest papers in this tradition, Eissa and Liebman (1996) evaluate the expansion of the EITC in the Tax Reform Act of 1986 and find positive and significant participation effects, but no effect on hours of work. Meyer and Rosenbaum (2001) exploit variation in the EITC up until 1996, controlling for changes to welfare (AFDC and food stamps), Medicaid, child care subsidies, and job training during this time period. Gelber and Mitchell (2012) exploit the same reform along with a large reform to the EITC in 1993 to examine the impact of taxes on the labor force participation of single women and their allocation of time to market work versus home production.

[5] Our sample omits married women and men. Rothstein (2010) points out that the wages of similarly skilled single and married women substantially diverged in the 1990s. For this reason, it seems reasonable to assume they operate in distinct labor markets. For men on the other hand, to the extent that they are substitutable for single women, we will be understating the size of each labor market and overstating the changes in market-level average tax rates. These effects will tend to work in opposite directions.

and year. The largely mirrors the definition of labor markets in (Rothstein, 2010). To generate micro-level variation in tax liabilities, we rely primarily on expansions to the federal EITC which differentially affected single women with and without kids. For macro-level variation, we rely on variation in state EITC levels, as well as variation in welfare benefits within states over time. To isolate purely exogenous variation in tax liabilities coming from policy reforms, we implement a simulated instruments approach similar to 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 women, is 0.72. This generally lines up with the range of estimates reported in the literature (Eissa et al., 2008). Focusing on subgroups, we find that the micro response is primarily driven by women with low levels of education, such as high school dropouts, with much smaller elasticities for bachelor degree holders. Our full sample estimate for the macro participation elasticity is 0.42. Across all subgroups, the macro effects tend to be smaller than the micro effects, although we emphasize that our macro estimates are imprecise. We also find that the employment responses to taxation is qualitatively similar to our participation responses. Finally, when we estimate (micro and macro) participation and employment responses using OLS, we generally find that the estimates are of an order of magnitude smaller than our IV estimates.

Finally, as an illustration, we use our empirical estimates to implement our sufficient statistics formula and calibrate the optimal income tax. Relative to the optimal tax schedule in Saez (2002), we find that if the macro participation response is less than the micro response, this moves it more towards an NIT-like tax schedule with a relatively larger lump sum payment to the non-employed combined with positive employment tax rates. We also show that calibrating our tax formula with smaller (employment and participation) macro responses has a much larger effect on the shape of the optimal tax profile (leading to a larger lump sum transfer and employment taxes), relative to calibrating the Saez (2002) formula with a smaller employment elasticity. This shows that it is misleading to simply calibrate existing tax formulas with macro employment elasticities, as standard intuition might suggest.

A number of recent papers have highlighted the distinction between micro and macro employment responses. The first paper to show that both are important for optimal policy is Landais et al. (2015), who consider a model of unemployment insurance (UI) with labor market spillovers and demonstrate that the optimal benefit level is a function of the gap between micro and macro unemployment duration elasticities. While our model is related in that it deals with spillover effects, the difference is that we consider multiple sectors of the labor market and focus on the optimal non-linear income tax; particularly, optimal transfers at the bottom of the income distribution. Landais et al. (2015) on the other hand have a single sector and focus on the optimal UI benefit level and how this should vary over the business cycle. Nevertheless, the distinction that the micro elasticity refers to responses that hold the job-finding rate (conditional on search intensity) and wages constant, while the macro elasticity allows the job-finding rate to adjust to UI benefits, is very similar to the distinction we introduce in our model. Partly inspired by Landais et al. (2015), some recent papers have tried to empirically estimate macro and micro effects of UI benefits (Lalive et al., 2013, e.g.) and job search assistance programs (Crépon et al., 2013, e.g.) on unemployment durations.[6]

The distinction between micro and macro responses has also played an important role in the recent empirical literature estimating extensive and intensive labor supply responses (See Chetty et al. (2011) and Chetty et al. (2012) for an overview). In this literature, the terms micro and macro responses are used slightly differently and instead characterize the source of variation in taxes used for identification. For macro, the source of variation is cross-country or business cycle whereas for micro, the source of variation is quasi-experimental. Differences between the two have been attributed to adjustment costs (Chetty et al., 2011) and optimization frictions (Chetty, 2012), an issue we abstract from in this paper. Instead, we consider responses that do (macro) or do not (micro) allow for certain equilibrium adjustment mechanisms.

This paper also relates to recent research on whether the generosity of UI benefits should depend on the state of the labor market. Unemployment benefits create a similar problem as traditional welfare benefits in that they provide transfers that are conditional on not working (or at least are at their maximum) and thus provide incentives not to work, while at the same time providing important insurance against hardship. Just as in the optimal taxation literature, the efficiency loss from providing UI is inversely related to the labor supply elasticities. Baily (1978), Chetty (2006), Schmieder et al. (2012), Kroft and Notowidigdo (2014) and Landais et al. (2015), derive welfare formulas where the marginal effect of increasing the generosity of unemployment benefits depends on the elasticity of unemployment durations with respect to the benefit generosity. These papers provide empirical evidence that the labor supply elasticities determining the optimal benefit durations (Schmieder et al., 2012) and levels (Kroft and Notowidigdo (2014) and Landais et al. (2015)) decline during periods of high unemployment and that the generosity of the UI system should therefore increase during these times. There are also papers that directly examine how labor supply responses to taxation vary with local labor market conditions. Closer to our setting, Herbst (2008) shows that the labor supply responses to a broad set of social policy reforms in the U.S. during the 1990s, such as EITC expansions, time limits, work requirements and Medicaid, are cyclical. Mogstad and Pronzato (2012) shows that labor supply responses to a "welfare to work"

[6] Crépon et al. (2013) evaluate an experiment of job placement assistance and find evidence of negative spillover effects (i.e., crowd-out onto untreated individuals). They find evidence that these spillover effects are larger when the labor market is slack and interpret this evidence as consistent with a model of job rationing (Landais et al., 2015). Lalive et al. (2013) show that the unemployment spells of individuals ineligible for UI were affected by a large expansion of Austria's UI benefits. Hagedorn et al. (2013) estimate large macro effects of unemployment insurance policies during the Great Recession. This is inconsistent with evidence that the micro effects of UI are small (Rothstein, 2011, Farber and Valletta, 2013). The authors stress the role of labor demand, although Marinescu (2014) does not find robust evidence of UI on vacancy creation.

reform in Norway are attenuated when the local unemployment rate is relatively high.

Finally, our work broadly relates to research which permit demand-side variables to determine employment outcomes and welfare participation for males and females. Blundell et al. (1987) shows that demand characteristics, such as unemployment rates, are important determinants of work for married females. Using the PSID, Ham and Reilly (2002) also find evidence that unemployment rates are significant predictors of work for males. While these papers focus on how demand-side factors affect the *level* of employment, our research explores whether such factors influence the *change* in employment in response to taxes and transfers. The role of demand side factors in affecting welfare use has been noted by others (see Hoynes (2000)), yet their normative implications have not been fully investigated so far.

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

II The theoretical model

In this section, we derive an optimal tax formula in a general model that is consistent with a rich set of labor market responses to taxation. Following Chetty (2009), we use this benchmark model to identify the sufficient statistics that are necessary to compute the optimal income tax. We do so first in the no-cross effect case where employment and participation responses are only on the extensive margin. This allows us to show the intuition of the main result before we go to the general formula that holds with arbitrary responses to taxes across labor markets. Our approach contrasts with papers that have incorporated unemployment into models of optimal taxation in a more structural way such as competitive models without unemployment (Mirrlees, 1971, Diamond, 1980, Saez, 2002), models with wage rigidity and job rationing (Lee and Saez, 2012) and matching models and Nash bargaining (Pissarides, 1985).[7] Below, we illustrate how these various structural models map into our sufficient statistic formula.

II.1 Setup

Labor markets

We start by generalizing the model in the appendix of Saez (2002) by introducing unemployment and wage responses to taxation. The size of the population is normalized to 1. There are $I+1$ "occupations" or income levels, indexed by $i \in \{0, 1, ..., I\}$. Occupation 0 corresponds to

[7]See Boadway and Tremblay (2013) for an excellent review of optimal income taxation in models with unemployment.

non-employment. All other occupations correspond to a specific labor market where the gross wage is w_i, the net wage (or consumption) is c_i and the tax liability is $T_i = w_i - c_i$. The assumption of a finite number of occupations is made for tractability. It is not restrictive as the case of a continuous wage distribution can be approximated by increasing the number I of occupations to infinity. The timing of our static model is:

1. The government chooses the tax policy.
2. Each individual m chooses the occupation $i \in \{0, ..., I\}$ to participate in. Individual heterogeneity only enters the model through the cost of search, as we indicate below.
3. For each labor market $i \in \{1, ..., I\}$, only 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 labor market and fail to find one and the *non-participants* who choose not to search for a job.[8] For each labor market $i \in \{1, ..., I\}$, k_i denotes the number of participants, $p_i \in (0, 1]$ denotes the fraction of them who find a job and are working, hereafter the *conditional employment probability*, and $h_i = k_i\, p_i$ denotes the number of employed workers. The number of unemployed individuals in labor market 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 verifies $h_0 = k_0 + \sum_{i=1}^{I} k_i(1 - p_i)$.

All the non-employed, whether non-participants or unemployed, receive the same welfare benefit denoted b.[9] 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 two defining the *employment*

[8] We simply assume job search intensity is either zero for non-participants or one for participants. Introducing a continuous job search intensity decisions as Landais et al. (2015) would add notational complexity while not substantially modifying the results.

[9] This is because the informational structure of our static model prevents benefits from being history-dependent. Moreover, as the government only observes income, it cannot distinguish non-participants from unemployed individuals. This latter assumption seems more realistic than the polar opposite one where the government can perfectly monitor job search. In this case, and if there is only one occupation, the government can provide full insurance to the unemployed.

tax.[10] 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.

Rather than specify the micro-foundations of the labor market, we use reduced-forms to describe the general equilibrium or *macro* responses of wages and conditional employment probabilities to tax policy $\mathbf{t}$. In labor market 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 by $p_i = \mathscr{P}_i(\mathbf{t})$. At this general stage, we are agnostic about the micro-foundations that lie behind these macro response functions and we only assume that these functions are differentiable, that $\mathscr{P}(\cdot)$ takes values in $(0,1]$ and that $0 < b < \mathscr{W}_1(\mathbf{t}) < ... < \mathscr{W}_I(\mathbf{t})$ for all tax policies $\mathbf{t}$. The latter assumption ensures that occupations indexed with a higher i correspond to labor markets with higher skills. The functions $\mathscr{W}_i(\cdot)$, $\mathscr{C}_i(\cdot)$ and $\mathscr{P}_i(\cdot)$ encapsulate all the effects of taxes, including those occurring through labor demand and wage setting responses.

Profits do not explicitly appear in the model. There are several justifications for this assumption. First, one can assume profits away. Competitive models with constant returns to scale or models with matching frictions on the labor market and free entry as (Mortensen and Pissarides, 1999) are models without profits. Alternatively, we can assume that occupation I actually corresponds to firm-owners for which w_I includes profits. The incidence of tax reforms on profits are then included as changes in w_I.[11]

Labor supply decisions

The structure of labor supply is as follows. We 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 an additional utility cost d_i for working in occupation i and a utility cost $\chi_i(m)$ for searching a job in labor market i.[12] Individual m thus enjoys a utility level equal to $u(c_i) - d_i - \chi_i(m)$ if she finds a job in labor market i, equal to $u(b) - \chi_i(m)$ if she is unemployed in labor market i, and $u(b)$ if she chooses not to search for a job. Let $\mathscr{U}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathscr{P}_i(\mathbf{t})\ (u(\mathscr{C}_i(\mathbf{t})) - d_i) + (1 - \mathscr{P}_i(\mathbf{t}))\, u(b)$ denote the *gross* expected utility of searching

[10] The literature uses instead the terminology *participation tax*, which we find confusing whenever unemployment is introduced. The *employment* tax $T_i + b$ captures the change in tax revenue for each additional *employed* worker. An additional *participant* being only employed with probability p_i, the change in tax revenue for each additional participant is only $(T_i + b)p_i$, which should correspond to the *participation* tax.

[11] In such a case, we need to assume full employment in occupation I. We also need to assume that for some individuals like capital owners or CEOs, only occupation I is available, while for others, only occupations $\{0, ..., I-1\}$ are available. As will be clear, our model allows for such restrictions on labor supply decisions.

Moreover, the budget constraint (1) is equivalent to a resources constraint. To see this, one simply needs to add $\sum_{i=1}^{I} c_i h_i$ on both sides of (1) to obtain $\sum_{i=1}^{I} w_i\ h_i = b\ h_0 + \sum_{i=1}^{I} c_i\ h_i + E$. The left-hand side, by adding labor income and profits, corresponds to total income in the economy, while the right-hand side corresponds to total expenditures in the economy. We rule out cases where profits are positive and distributed to all individuals.

[12] We denote $\chi_0(m) = 0$. We furthermore assume that $\chi_i(m) = +\infty$ if individual m does not have the required skill to work in occupation i.

for a job in occupation i, absent any participation cost, as a function of the tax policy $\mathbf{t}$, and let U_i denote its realization at a particular point of the tax system.[13] Let $U_0 = u(b)$ be the utility expected out of the labor force.

Individual m expects utility $U_i - \chi_i(m)$ by searching for a job in labor market i. She chooses to search in labor market i if and only if $U_i - \chi_i(m) > U_j - \chi_j(m)$ for all $j \in \{0, ..., I\} \setminus \{i\}$. The set of individuals choosing to participate in labor market i is therefore $M_i(U_1, ..., U_I, u(b)) \stackrel{\text{def}}{\equiv} \{m | i = \arg\max_{j \in \{0,...,I\}} U_j - \chi_j(m)\}$. Assuming that participation costs $(\chi_1, ..., \chi_I)$ are distributed in the population in a sufficiently smooth way and denoting $\mu(.)$ the distribution of individuals, the number k_i of participants in labor market i is a continuously differentiable function of expected utility in each occupation through: $k_i = \hat{\mathcal{K}}_i(U_1, ..., U_I, u(b)) \stackrel{\text{def}}{\equiv} \mu(M_i(U_1, ..., U_I, u(b)))$. Participation decisions are determined through:

$$k_i \equiv \mathcal{K}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \hat{\mathcal{K}}\left(\mathscr{U}_1(\mathbf{t}), ..., \mathscr{U}_I(\mathbf{t}), u(b)\right) \tag{2}$$

Finally, employment is given by:

$$h_i = \mathcal{H}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathcal{K}_i(\mathbf{t}) \mathscr{P}_i(\mathbf{t}) \tag{3}$$

Micro vs. Macro Responses

A crucial distinction is the difference between macro and micro participation responses to taxes. We define the *micro* participation response 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 instance, the case for tax reforms frequently considered in the micro-econometric literature that affect only a small subset of the population, so that the general equilibrium effects of the reform on wages and conditional employment probabilities can be safely ignored. The micro response of expected utility is thus $-p_i u'(c_i)$. Moreover, from Equation (2), as taxes affect participation decisions only through expected utility levels in each occupation, the micro participation response is given by:

$$\left.\frac{\partial \mathcal{K}_i}{\partial T_j}\right|^{\text{Micro}} \stackrel{\text{def}}{\equiv} -p_j u'(c_j) \frac{\partial \hat{K}_i}{\partial U_j} \tag{4}$$

Conversely *macro* responses encapsulates wage and conditional employment probability responses. The macro response of expected utility is therefore:

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

The term within brackets on the right-hand side of (5) in particular describes how the wage and conditional employment probability responses induce a gap between macro and micro expected

[13] U_i is identical across all participants because the conditional employment probability p_i and the wage w_i are identical across participants in labor market i and in particular do not vary with $(\chi_1(m), ..., \chi_I(m))$.

utility responses. Using (2) and (5), the macro participation response is given by:

$$\frac{\partial \mathcal{K}_i}{\partial T_j} = \sum_{\ell=1}^{I} \frac{\partial \mathcal{U}_\ell}{\partial T_j} \frac{\partial \hat{\mathcal{K}}_i}{\partial U_\ell} = \sum_{\ell=1}^{I} \left[\frac{\partial \mathcal{C}_\ell}{\partial T_j} + \frac{\partial \mathcal{P}_\ell}{\partial T_j} \frac{u(c_\ell) - d_\ell - u(b)}{p_\ell \, u'(c_\ell)} \right] p_\ell \, u'(c_\ell) \, \frac{\partial \hat{\mathcal{K}}_i}{\partial U_\ell} \tag{6}$$

The micro and macro participation responses differ for two main reasons. First, utility levels in the occupation that experiences the tax change can be affected by change in the wage and in the conditional employment probability in that occupation, as we demonstrate below. For micro responses, gross wages are held constant, thus $\frac{\partial \mathcal{C}_j}{\partial T_j} = -1$ and taxes are passed through one for one to the worker, while employment probabilities are also fixed and thus $\frac{\partial \mathcal{P}_j}{\partial T_j} = 0$. For macro responses on the other hand, tax adjustments may affect gross wages in a variety of ways $\frac{\partial \mathcal{C}_j}{\partial T_j} \neq -1$ while employment probabilities may also change $\frac{\partial \mathcal{P}_j}{\partial T_j} \neq 0$, e.g. due to changes in labor supply in that occupation or due to changes in vacancy creation by employers, as we demonstrate more formally below. Second, utility levels can also be affected by change in the tax liability in other occupations, explaining the summation over all occupations in (6). This could be for example because increasing taxes in occupation j may lead firms to adjust their composition of labor inputs and may change labor demand for other occupations. Moreover, it may be because the workers who are less likely to search for jobs in occupation j may look for jobs in other occupations which will thus change equilibrium outcomes in those occupations.

Social objective

We assume that the government maximizes a weighted utilitarian welfare objective that depends only on individuals' expected utilities:

$$\Omega(U_1, ..., U_I, u(b)) = \int \gamma(m) \left(\max_i \quad U_i - \chi_i(m) \right) \, d\mu(m) \tag{7}$$

where the weights $\gamma(m)$ may vary across individuals. In the particular case where the utility function $u(\cdot)$ is linear, it is the variation of weights with the characteristics of individuals through the heterogeneity in $\gamma(\cdot)$ that generates the social desire for redistribution, while if individual utility is concave the desire for redistribution comes (also) from individual risk aversion.[14]

The optimal policy

The government chooses the tax policy $\mathbf{t} = (T_1, ..., T_I, b)'$ to maximize (7) subject to the budget constraint (1). Let $\lambda > 0$ denote the Lagrange multiplier associated with the latter constraint.

[14] It is straightforward - and does not change our results below - to generalize this social welfare function to the case where the social planners maximizes an arbitrary concave function of individual expected utilities integrated over the population.

Following Saez (2001, 2002), we define the marginal social welfare weight of workers in occupation $i \in \{1, ..., I\}$ as:

$$g_i \stackrel{\text{def}}{\equiv} \frac{1}{k_i} \frac{\partial \Omega}{\partial U_i} \frac{u'(c_i)}{\lambda} = \frac{p_i\, u'(c_i) \int_{m \in M_i} \gamma(m)\, d\mu(m)}{\lambda\, h_i} \tag{8}$$

The social weight g_i represents the social value in monetary terms of transferring an additional dollar to an individual working in occupation i. It captures the *micro* effect on the social objective of a unit decrease in tax liability, expressed in monetary terms. Absent wages and conditional employment probabilities responses, the government is indifferent between giving one more dollar to an individual employed in labor market i and g_i more dollars of public funds. Using Equations (5) and (8), we get the following lemma.

Lemma 1. *The first-order condition with respect to the tax liability T_j in labor market j is:*

$$0 = \underbrace{h_j}_{\text{Mechanical effect}} + \underbrace{\sum_{i=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_j} (T_i + b)}_{\text{Behavioral effects}} + \underbrace{\sum_{i=1}^{I} \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{p_i\, u'(c_i)} \right] g_i\, h_i}_{\text{Social Welfare effects}} \tag{9}$$

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 number h_j of employed individuals in occupation j.

2. **Behavioral effects**: A unit increase in T_j induces a change $\partial \mathcal{H}_i / \partial T_j$ in the level of employment in occupation i. 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 affects the social objective in monetary terms and using Equations (5) and (8), we get that the social welfare effect of tax T_j in occupation i is: $g_i h_i \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{p_i\, u'(c_i)} \right]$. Note that because the social welfare function depends on expected utility U_i, the labor supply responses only modifies 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. This is the reason why these "market spillovers" show up in the social welfare effect through the term within brackets, unlike the participation responses. Because the social objective as well as participation decision depend on the tax policy only through expected utility levels in each occupation, the same terms $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{p_i\, u'(c_i)}$ describe how

macro social welfare effects differ from micro ones and how macro participation responses differ from micro ones.

Optimal benefit level

Finally, for the sake of completeness, the first-order condition with respect to the welfare benefit b is (see Appendix A.1):

$$0 = -h_0 + \sum_{i=1}^{I} (T_i + b) \frac{\partial \mathcal{H}_i}{\partial b} + g_0 h_0 + \sum_{i=1}^{I} g_i h_i \left[\frac{\partial \mathscr{C}_i}{\partial b} + \frac{1}{p_i} \frac{\partial \mathscr{P}_i}{\partial b} \frac{u(c_i) - d_i - u(b)}{u'(c_i)} \right] \tag{10}$$

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

$$g_0 \overset{\text{def}}{\equiv} \frac{u'(b)}{h_0} \left[\int_{m \in M_0} \frac{\gamma(m)}{\lambda} d\mu(m) + \sum_{i=1}^{I} \frac{g_i}{u'(c_i)} k_i (1 - p_i) \right] \tag{11}$$

In particular, if we furthermore assume there is no income effects so that $\sum_{i=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial b}$, $\sum_{i=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_i} = \frac{\partial \mathscr{P}_i}{\partial b}$ and $\sum_{i=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_i} = \frac{\partial \mathcal{H}_i}{\partial b}$, we get that the weighted sum of social welfare weights is 1 (See Appendix A.1):

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

II.2 The sufficient statistics optimal tax formula

To numerically implement the optimal tax formula in equation (9), one must know the gap in utilities between employment and non-employment, the responses of net wages to taxation $\frac{\partial \mathscr{C}_i}{\partial T_j}$ and the responses of the conditional employment probabilities to taxation $\frac{\partial \mathscr{P}_i}{\partial T_j}$ that appear in the social welfare effects. We now show that there is a simpler representation for the optimal tax formula (9) in terms of the *macro* $\frac{\partial \mathcal{K}_i}{\partial T_j}$ and *micro* participation responses $\left.\frac{\partial \mathcal{K}_i}{\partial T_j}\right|^{\text{Micro}}$. The advantage of this representation is that we may apply conventional econometric techniques to estimate these terms.

The no-cross-effect case

To simplify the exposition and develop intuition, we begin with the "no-cross-effect" case where we assume for simplicity that $\partial \mathscr{W}_i / \partial T_j = \partial \mathscr{C}_i / \partial T_j = \partial \mathscr{P}_i / \partial T_j = \partial \hat{\mathcal{K}}_i / \partial U_j = 0$ for $i \neq j$ and $i \neq 0$. The last equality implies that labor supply responses are concentrated along the extensive margin. Moreover, we get from (5) that $\partial \mathscr{U}_i / \partial T_j = 0$, which together with (2) and (3) imply that: $\partial \mathcal{K}_i / \partial T_j = \partial \mathcal{H}_i / \partial T_j = 0$ 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. The no-cross-effect environment includes the model of Landais et al. (2015) where the wage depends on the level of tax liability but not on the marginal tax rate.

In the no-cross effect case, Equations (4) and (6) imply that we may express the macro participation response in terms of the micro participation response in the following way:

$$\frac{\partial \mathcal{K}_j}{\partial T_j} = \left[\frac{\partial \mathscr{C}_j}{\partial T_j} + \frac{\partial \mathscr{P}_j}{\partial T_j}\frac{u(c_j) - d_j - u(b)}{p_j\, u'(c_j)}\right] \left.\frac{\partial \mathcal{K}_j}{\partial T_j}\right|^{\text{Micro}}$$

The formula (9) for the optimal tax liability in occupation j then simplifies to:

$$0 = h_j + \frac{\partial \mathcal{H}_j}{\partial T_j}\,(T_j + b) + \frac{\frac{\partial \mathcal{K}_j}{\partial T_j}}{\left.\frac{\partial \mathcal{K}_j}{\partial T_j}\right|^{\text{Micro}}} g_j h_j \tag{12}$$

To better relate this expression to the optimal tax literature, we define the micro participation elasticity as $\pi_j^m \overset{\text{def}}{\equiv} -\frac{c_j - b}{k_j}\left.\frac{\partial \mathcal{K}_j}{\partial T_j}\right|^{\text{Micro}}$. This elasticity measures the percentage of employed workers in i who leave the labor force when the tax liability is increased by 1 percent, holding wages and the conditional employment probabilities fixed. Next, we define the macro employment elasticity as $\eta_j \overset{\text{def}}{\equiv} -\frac{c_j - b}{h_j}\frac{\partial \mathcal{H}_j}{\partial T_j}$. From (3), the macro *employment* response η_j verifies $\eta_j = \frac{c_j - b}{p_j}\frac{\partial \mathscr{P}_j}{\partial T_j} + \pi_j$. In particular, it encapsulates 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. Moreover, wage and unemployment responses modify the macro participation responses π_j from the micro ones π_j^m, as discussed above.

Proposition 1. *The optimal tax formula 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 compared to pure extensive case without unemployment case considered by Diamond (1980), Saez (2002) and Choné and Laroque (2005, 2011) where it is: $\frac{T_j + b}{c_j - b} = \frac{1 - g_j}{\pi_j^m}$.

There are two key differences between Equation (13) and Equation (4) in Saez (2002). First, the denominator in (13) corresponds to the macro employment elasticity, whereas in Saez (2002) it is the micro participation elasticity that appears. Second, equation (13) modifies the social marginal welfare weight by the ratio of the macro to micro participation elasticity. The response of expected utility may be different at the macro and micro levels. This is because the macro responses encapsulate not only the direct effect of a tax change on consumption, but also the indirect effects of a tax change on the wage $\frac{\partial \mathscr{W}_i}{\partial T_i} \neq 0$ and on the conditional employment probability $\frac{\partial \mathscr{P}_i}{\partial T_i} \neq 0$. The ratio between the micro and macro expected utility responses corresponds exactly to the ratio of the macro to the micro participation elasticities. So the welfare effect may be larger or lower than the social welfare weight g_i. To understand why, consider a decrease in tax liability T_j. 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 decreases in the gross wage when $\frac{\partial \mathscr{W}_j}{\partial T_j} > 0$, so the responses of wage attenuates the direct impact on social welfare. Finally, the decrease in tax liability also typically triggers a rise in job creation, i.e. $\frac{\partial \mathscr{P}_j}{\partial T_j} < 0$, 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 responses dominates (is dominated by) the impact of the wage responses. In particular, if the effect of the tax on the conditional employment probability happens only though a labor demand response, the macro participation response is higher than micro one if the labor demand elasticity is high enough. We therefore get:

Corollary 1. *In the no-cross-effect case, the optimal employment tax 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 micro over macro participation elasticity, instead of one without unemployment and wage responses, a condition that can be easily tested.

The case with cross effects

We now turn back to the general formula with cross effects, where matrix notation turns out to be convenient. For $f = \mathcal{K}, \hat{\mathcal{K}}, \mathcal{H}, \mathscr{U}, \mathscr{P}, \mathscr{W}$ and $x = T, U$, we denote $\frac{\mathbf{df}}{\mathbf{dx}}$ the square matrix of rank I whose term in row j and column i is $\frac{\partial f_i}{\partial x_j}$ for $i, j \in \{1, ..., I\}$.[15] Symmetrically, the matrix of micro responses are denoted $\left.\frac{\mathbf{df}}{\mathbf{dx}}\right|^{\text{Micro}}$. Moreover, $\mathbf{h} = (h_1, ..., h_I)'$ denotes the vector of employment levels, $\mathbf{g}\,\mathbf{h} = (g_1 h_1, ..., g_I h_I)'$ denotes the vector of welfare weights times employment levels and $\cdot$ denotes the matrix product. Appendix A.2 then shows that market spillover terms $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{p_i \; u'(c_i)}$ that appear in the social welfare effects in the optimal tax formula (9) still correspond to the ratio of macro over micro participation responses. The only difference is that in the presence of cross effects, this ratio should be understood in matrix terms. We thus get the following generalization of the optimal tax (12) in the presence of cross effects:

Proposition 2. *If* $\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{Micro}$ *is invertible, the optimal tax system for occupations* $i = \{1, ..., I\}$ *solves the following system of equations in matrix form:*

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) - \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{Micro} \right)^{-1} \cdot (\mathbf{g}\,\mathbf{h}) \tag{14}$$

[15] In particular, these matrices do not include partial derivatives with respect to b, nor do they include partial derivatives for occupation 0.

Equation (14) is expressed in terms of sufficient statistics. It implies that the ratio (in matrix terms) of macro to micro participation responses are the sufficient statistics to estimate, instead of the market spillover terms that depend on net wage $\frac{\partial \mathscr{C}_i}{\partial T_j}$ and conditional employment probability responses $\frac{\partial \mathscr{P}_i}{\partial T_j}$. Intuitively, because the social welfare function is assumed to depend only on expected utilities, the market spillovers that appear in the social welfare effects in (9) coincide with the terms $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j} \frac{u(c_i)-d_i-u(b)}{p_i u'(c_i)}$ that describe how the macro responses of expected utility differ from the micro ones (see (5)). Moreover, because participation decisions depend only on expected utility as well, these market spillovers are entirely captured by the matrix ratio of macro over micro participation responses. Importantly, the gap between micro and macro responses does not matter for the behavioral effects, but only for the social welfare effects. This is because the matrix $\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}$ of macro employment responses already encapsulates the unemployment and wage responses in addition to the micro participation responses.

II.3 The links between the optimal tax formula and micro-foundations of the labor market

The key result of our paper is that the optimal tax schedule can be implemented using the macro employment responses as well as the ratio of the macro over micro participation responses. While the advantage of our approach is that it does not rely on any assumptions about the structure of the labor market, the downside is that it does not provide any guidance for how to interpret differences between macro and micro participation responses. In this section, we discuss how different micro-foundations yield different predictions for the relative magnitude of micro and macro participation (and to a lesser degree employment) responses. This serves to build intuition for the macro-micro gap and thereby what economic forces push the optimal tax at the bottom towards an EITC or NIT, while at the same time highlighting how our framework encompasses standard models of the labor market. Since providing an exhaustive exposition of possible models would go beyond the scope of the paper, we instead focus on two main paradigms: search-matching models and job-rationing models. These two classes of models provide a determination for the level of employment and for the unemployment rate as a function of the wage and possibly labor supply. It is worth noting that the different micro-foundations for the gross wage, such as a minimum wage, an efficiency wage or a bargained wage can be included in either of these two frameworks. We start with the search-matching paradigm before presenting the job-rationing paradigm. We then briefly discuss the competitive model and finally models with a wage moderating effect of tax progressivity.

Search and matching models with constant returns to scale (CRS)

In its simplest version, the search-matching framework (Diamond, 1982, Pissarides, 1985, Mortensen and Pissarides, 1999, Pissarides, 2000) assumes a linear (constant returns to scale) production function and a matching function which gives the number of jobs created as a function of the number of vacancies and the number of job seekers. Firms employ more workers the lower the gross wage (which makes more rewarding rewarding for firms to hire a worker) and the more numerous job-seekers there are (which decrease the search congestions from firms' viewpoint thereby easing their recruitment). In the model, the conditional employment probability p_i is a decreasing function $\mathscr{L}_i(\cdot)$ of the gross wage and is independent of the number of job-seekers.[16] Therefore, a policy reform that increases labor supply, without affecting the gross wage, leads to a rise in employment in the same proportion as the rise in labor supply, but does not affect the employment probability.

If we consider a version of the matching model where wages are fixed, than the conditional employment probabilities are fixed, so the macro participation responses are equal to the micro ones. If we instead consider a version of the matching model where wage setting is based on wage bargaining, taxes may affect the outside option for workers as well as the match surplus and thus equilibrium wages and in turn conditional employment probabilities. To build intuition, consider the case with risk neutral workers (hence $u(c) \equiv c$) and proportional bargaining. In such a setting, workers receive an exogenous share $\beta_i \in (0,1)$ of the total match surplus $y_i - T_i - d_i - b$, so the wage is given by:[17]

$$w_i = \mathscr{W}_i(T_i, b) \equiv \beta_i \, y_i + (1 - \beta_i)(T_i + d_i + b) \tag{15}$$

Combining the labor demand relation $p_i = \mathscr{L}_i(w_i)$ with the wage equation (15) and the assumption that labor supply responses are concentrated along the extensive margin provides a complete search-matching micro-foundation for the no-cross effect economy. The following proposition shows that the macro-micro participation gap is directly linked to the bargaining weights and the elasticity of the matching function with respect to the number of job-seekers $\mu_i \in (0,1)$:

Proposition 3. *In the search-matching economy with proportional bargaining* (15)*, the micro and macro participation responses are equal either when the workers have full bargaining power so there is no wage responses, or when the Hosios (1990) condition $\beta_i = \mu_i$ for a decentralized economy without tax and transfer to be socially efficient is verified. If $\beta_i < \mu_i$ the macro response is lower then micro one. If $\mu_i < \beta_i < 1$ the macro response is larger then micro one.*

An increase in tax liability has three effects on expected utility, thereby on participation decisions. First, absent wage and conditional employment response, a rise in T_i has a *direct* negative

[16] We derive in AppendixA.3 this standard result, as well as the proof of Proposition 3 below.

[17] A similar expression for wage bargaining appears in Jacquet et al. (2014) and in Landais et al. (2015).

impact at the micro level (holding w_i and p_i constant) as it reduces the net wage and thus incentives to work and to participate. Second, at the macro level, gross wages increases (through bargaining) attenuating the direct labor supply effect. Finally, the gross wage increase triggers a reduction in labor demand that amplifies the direct effect at the macro level. If the workers get all of the surplus (i.e. if $\beta_i = 1$), wages do not respond to taxation ($\frac{\partial \mathscr{W}_i}{\partial T_i} = 0$), the conditional employment probabilities are not affected so the micro and macro responses to participation are identical. On the other hand, if $\beta_i < 1$, the conditional employment probability effect dominates (is dominated by) the wage effect whenever the labor demand elasticity is (not) sufficiently elastic, which happens when the matching elasticity μ_i is higher (lower) than the bargaining share β_i. Propositions 1 and 3 imply that the optimal employment tax rate on the working poor is more likely to be negative in the no-cross effect DMP case than in the pure extensive case if the workers' bargaining power is inefficiently high, i.e, is higher than the bargaining power prescribed by the Hosios (1990) condition.[18] Therefore, in the DMP model the macro micro participation gap can be higher or lower than one, attenuating or reinforcing the arguments in favor of a negative participation tax at the bottom.

Job-rationing models

An older tradition in economics has proposed job rationing to explain unemployment. In contrast to the matching framework, the job-rationing framework assumes search frictions away and considers that each type of labor exhibits decreasing marginal productivity. In each labor market, employment is determined by the equality between the marginal product and the wage. Unemployment occurs whenever the wage is set above its market-clearing level. This theory of unemployment that Keynes (1936) attributed to Pigou was formalized in the disequilibrium theory (Barro and Grossman, 1971) and further developed by making wages endogenously above the market clearing level (McDonald and Solow, 1981, Shapiro and Stiglitz, 1984, Akerlof and Yellen, 1990).[19]

To develop some intuition about the macro-micro participation gap in job-rationing models, we now consider a model with a single type of labor that exhibits a decreasing marginal productivity and a fixed gross wage w. This can occur for instance as a result of a minimum wage regulation. The fixed wage determines the level of employment h, independently of the number of participants.[20] We assume that individuals who participate face a heterogeneous participation

[18]As $\frac{\pi_j}{\pi_j^m} = \frac{\beta_j}{\mu_j}$ from (23), Equation (13) becomes $\frac{T_j+b}{c_j-b} = \frac{1-\frac{\beta_j}{\mu_j} g_j}{\eta_j}$ which corresponds to (19b) in Jacquet et al. (2014).

[19]The Keynesian and New Keynesian theories of unemployment in addition assume nominal rigidities to give a transitional role to aggregate demand management policies. See also Michaillat and Saez (2015) for an extension of the new Keynesian model in which disequilibrium due to price rigidity are smoothed by matching functions on both the labor and the product market.

[20]Note that with a fixed wage, it is no longer equivalent whether the firm or the worker pays the tax. If the firm pays

cost χ that is sunk upon participation. The k participants face the same probability $p = h/k$ to be employed, whatever the participation cost χ they incur if they participate. In such a framework, a tax cut in T triggers a rise in participation at the micro level. However, provided that this tax cut occurs for a fixed wage, employment does not change, so the macro employment response is nil. Therefore, as the number of participants increases, the probability to be employed is reduced, which attenuates the participation responses at the macro level, as compared to the micro one. As a result, the optimal employment tax on the working poor is more likely to be positive in this job-rationing model without cross effect than in the pure extensive case.

There are different job-rationing models in the literature. For instance, in Lee and Saez (2012), there are different types of labor that are perfect substitutes, the minimum wage policy is explicitly an additional policy instrument and efficient rationing is assumed, so that the probability to be employed varies across participants as a function of their private cost upon working. Wages can also be made endogenous through union bargaining (McDonald and Solow, 1981) or through efficiency wages (Shapiro and Stiglitz, 1984, Akerlof and Yellen, 1990). Job rationing can also be analyzed within a search-matching framework if decreasing returns to scale is assumed for the production function, as in Michaillat (2012). As in a job-rationing model without matching, the macro employment effect would be dampened compared to the micro one and conditional employment probabilities would fall in response to a tax decrease. This in turn generates a gap in the micro and macro participation response that captures the spillover effect on the labor market. While decreasing returns to scale may not be realistic in the long run, it may be plausible at least in the short-run during recessions with aggregate demand shortfalls. (Landais et al., 2015) discuss this possibility as a possible reason that the effect of unemployment insurance benefits on employment may be larger when the labor market is tight than when it is slack and thus the moral hazard associated with UI may be less severe during a crisis. For the same reason it may be that reductions in tax levels may have a larger effect on employment in recessions than in booms and the optimal policy during recessions may look more like an NIT.[21]

Competitive models

Like job-rationing models, competitive models assume search frictions away. However, these models assume that in each labor market, the gross wage adjusts to clear the labor market so there is no unemployment. If, in addition, the technology exhibit constant returns to scale and perfect substitution across the different types of labor, labor demand is perfectly elastic and our model

the tax, then a tax cut reduces the cost of labor and increases labor demand. In this case, the government controls not only the total tax liability in an occupation, but also the cost of labor and thereby the employment level. Lee and Saez (2012) provides conditions where the government finds optimal to set the cost of labor above the market-clearing level, thereby generating unemployment in a job-rationing model.

[21] Though note that we have a static framework which may not be well suited to determine time-varying optimal taxes over the business cycle.

reduces immediately to Saez (2002).[22] In such a model, there is no difference between macro and micro responses, so the optimal tax formula depends only on the macro (or micro) employment effect of taxes. On the other hand, consider a competitive model with constant returns to scale technology and flexible wages: there would be no unemployment, but wages may adjust to taxes due to imperfect substitution across the different types of labor. In this case the micro and macro employment responses may be different due to the wage adjustments in each labor market, but the participation gap would still capture these spillover effects. Saez (2004) showed that in such a model the optimal tax formula can be expressed using only the micro employment response and in that case takes on the same form as Saez (2002). In Appendix A.4, we show that this result remains valid if unemployment rates are positive but exogenous. So, the optimal employment tax is negative when the social marginal welfare weight exceeds one. However, even in this case, our optimal tax formula (14) remains valid.

Wage moderating effects of tax progressivity

Another strand in the literature has stressed the possibility that increases in tax progressivity may actually increase employment. For example in the monopoly union model, unions set the wage to maximize the expected utility of its members, which is increasing in the net wage and in the level of employment. Since the level of employment is decreasing in the gross wage, unions do not want to push the wage too high. If tax rates increase (become more progressive) the wedge between net and gross wages increases and therefore a one unit increase in the net wage will have to be traded off against a larger loss in employment. Thus unions may actually accept a lower gross wage in response to an increase in tax progressivity and tax increases may increase employment.[23] The main consequence of introducing the wage moderating effect of tax progressivity into the model is to make the matrix $\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}$ and therefore the matrices $\frac{\mathbf{d}\mathscr{P}}{\mathbf{dT}}$, $\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}$, $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}$ and $\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}$ non-diagonal. The wage moderating effect of tax progressivity is therefore an argument against the no-cross effect restriction, which is different than the presence of labor supply responses along the intensive margin.[24]

[22] Assuming fixed w_i and p_i, equation (9) collapses to the optimal tax formula (11) in the Appendix of Saez (2002). This formula can be further specialized by assuming that labor supply responses are concentrated along the intensive margin (Mirrlees (1971) and Saez (2002, Equation (6))), along the extensive margin (Diamond (1980), Saez (2002, Equation (4)) and Choné and Laroque (2005, 2011)) or both (Saez (2002, Equation (8)))

[23] This result has been obtained in a Monopoly unions model with job rationing by Hersoug (1984), in a union bargaining model by Lockwood and Manning (1993). A very similar result can also hold in the efficiency wage model of Pisauro (1991) or within the matching framework with Nash bargaining (Pissarides, 1985, 1998), or with the bargaining model of top income earners of Piketty et al. (2014). Evidence for this wage moderating effect of tax progressivity can be found in Malcomson and Sartor (1987), Holmlund and Kolm (1995), Hansen et al. (2000) and Brunello and Sonedda (2007), while Manning (1993) and Lehmann et al. (2014) provide some empirical support for the unemployment reducing effect of tax progressivity.

[24] In the context of our framework reduced to the case with two occupations $I = 2$, these models imply that the wage functions $\mathscr{W}_i$ not only verify $\frac{\partial \mathscr{W}_2}{\partial T_2} > 0$ and $\frac{\partial \mathscr{W}_1}{\partial T_1} > 0$, as in the proportional bargaining case, but also that the marginal tax rate, as approximated by $T_2 - T_1$, has a wage moderating and unemployment reducing effect. This implies

III Estimating Sufficient Statistics

In this section, we estimate the sufficient statistics necessary to implement our optimal tax formula, namely the macro employment response to taxes, and the micro and macro participation responses. We follow the large empirical literature on the effects of the EITC and welfare reform in the U.S. and focus on the employment and participation responses of single women 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. These policy reforms provide sufficient variation to identify both micro and macro participation responses and macro employment responses.

III.1 Data

Current Population Survey (CPS)

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 women age 18 to 55 who are not in the military or enrolled full time in school, college or university. 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). [25]

Panel A of Table 1 shows descriptive statistics for the demographic characteristics of single women in the March CPS by educational attainment, pooling all years from 1984 to 2011.[26] The age range is pretty similar across the four education groups but there are large differences in the distribution of number of children, with lower educated single women being much more likely to be mothers. This is likely due to our sample restriction to single women since higher educated mothers are more likely to be married. Additionally, low educated women are more likely to be

that $\frac{\partial \mathscr{W}_2}{\partial T_1} > 0 > \frac{\partial \mathscr{W}_1}{\partial T_2}$. Within a matching model, using $p_i = \mathscr{L}_i(w_i)$, we obtain $\frac{\partial \mathscr{P}_2}{\partial T_2} < 0$ and $\frac{\partial \mathscr{P}_1}{\partial T_1} < 0$, but also $\frac{\partial \mathscr{P}_2}{\partial T_1} < 0 < \frac{\partial \mathscr{P}_1}{\partial T_2}$. Hence, making the tax schedule more progressive by increasing T_2 and decreasing T_1 increases employment in both occupations, which the government finds beneficial whenever employment taxes remain positive. Hence, compared to the proportional bargaining case, the case with a wage moderating/unemployment reducing effect of tax progressivity leads to a more progressive optimal tax schedule as formally shown by Hungerbühler et al. (2006), Lehmann et al. (2011).

[25] For complete details on sample construction and variable definitions, please see the online Appendix.

[26] 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.

black or hispanic than high educated ones. Panel B displays labor market variables by educational attainment. Lower educated women are much less likely to be in the labor force than higher educated ones and also experience much 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 build a calculator that computes taxes and transfers at (nominal) income levels for single women, depending on the number of children, state and year. We assume that a woman is filing as the head of the household and claims 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.[27] 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 also 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.[28] Rather than incorporate all of these policies explicitly into our empirical framework, we multiply benefits by take-up rates constructed from the Survey of Income and Program Participation (SIPP). The new eligibility requirements are reflected in lower observed take-up 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

[27]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.

[28]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.

impute an individual's tax liability following the approach taken in Eissa and Hoynes (2004) and Gelber and Mitchell (2012). We begin by running separate regressions for each education group (e) and year (t) of log annual earnings for individual m on state fixed effects ($\delta_{e,s,t}$) and control variables ($X_{m,e,s,t}$): [29]

$$log(w_{m,e,s,t}) = \delta_{e,s,t} + X_{m,e,s,t}\pi_{e,t} + \epsilon_{m,e,s,t} \tag{16}$$

The control variables include state fixed effects, 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.[30] We then use predicted earnings to impute an individual's tax liability using TAXSIM and the benefit calculator described above.

Below 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. To address this concern, we also rely on a simulated instrument approach Currie and Gruber (1996).[31] This approach isolates policy variation in tax liabilities, and holds fixed the distribution of income and demographics during the sample period.

There are several steps that we take to implement this procedure. To construct the simulated *micro* tax liabilities, we first compute real earnings in 2010 dollars for each individual in the sample. Second, using earnings for the full sample of individuals across all years 1984-2011, we construct the percentiles of the empirical earnings distribution. Third, we compute for each education group, the percentage of workers that fall into each centile. Fourth, for each year, we compute the nominal earnings level in each centile, conditional on real earnings in that year being within the bounds of the centile from step 2. Fifth, for each year, we take the nominal earnings level in each centile and we compute tax liabilities separately by number of children for each state, using the tax and transfer calculator. In the last step, for each education group, year, state and number of children, we compute the weighted-mean of the tax liabilities across centiles using the (time- and state-invariant) education distribution from step 3 as weights. This leaves us with instruments that are cell means, where the cells are defined by education group, year, state, and number

[29]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.

[30]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 bachelor degree holders seemed too low.

[31]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.

of children, with variation driven solely by exogenous changes in the tax code, and not by endogenous changes in the earnings distribution. Finally, for the simulated *macro* tax liability, we aggregate micro tax liabilities across family types using weights for number of children that vary by education group, but are time- and state-invariant. All tax liabilities are adjusted for inflation using the consumer price index for all urban consumers with 2010 as the base year. The simulated cell average (micro and macro) tax liabilities are then matched back to the CPS data and used as instruments for imputed tax liabilities, among individuals in a given cell, in a two-stage least squares regression.

Panel C of Table 1 shows the mean 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. 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, 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 take-up rates which are much less than 100 percent during our sample period. The last four rows report take-up rates, as estimated in the SIPP. Each individual in the CPS is assigned a take-up rate that we calculate from the SIPP based on education, income and year. The table reports the average of the assigned take-up rates separately for AFDC/TANF and food stamps, and also pre- and post-1996. We see that for high school dropouts, take-up rates are roughly 50 percent for AFDC/TANF but fall to 20 percent post-1996. For food stamps, take-up rates are much more comparable pre- and post-1996 and equal to roughly 40 percent.[32] These take-up rates decrease with education which reflects diminishing eligibility as earnings increase.

III.2 Empirical Method

Specification of Labor Markets

In the theoretical model, individuals are separated into I distinct labor markets or income groups. For our empirical analysis, a key difficulty is ranking individuals, including the non-employed, according to their potential income if they work. For this purpose, we approximate the labor market an individual may participate in by her educational attainment (high school dropout, high school graduate, some college, bachelor), state and year. We assume that individuals are perfect substitutes within labor markets and use (e, s, t) to denote these cells. This labor market definition is consistent with Rothstein (2010). We also adopt a pooling assumption which allows the partial derivatives of taxes on participation and employment to vary by education, but not by

[32]For AFDC/TANF, we calculate take-up rates based on sample of mothers since single women with no children are not eligible for these programs.

state and year.

Estimating Micro and Macro Participation Responses and Macro Employment Responses

Equation (14) shows that the optimal tax schedule can be expressed in terms of (the matrix) of macro employment responses and the ratio (in matrix terms) of macro to micro participation responses. Ideally one would attempt to estimate the matrix of macro participation responses $\frac{\partial \mathcal{K}_i}{\partial T_j}$, the matrix of micro participation responses $\frac{\partial \mathcal{K}_i}{\partial T_j}\Big|^{\text{Micro}}$ and the matrix of macro employment responses $\frac{\partial \mathcal{H}_i}{\partial T_j}$ for all labor markets i, j. However, this would lead to a very large number of cross effects to estimate that would be difficult to identify, especially the macro responses. Thus, for the purpose of estimation, we focus on the no cross effects case where the above mentioned matrices are diagonal.

To obtain an econometric specification for the participation decision that is motivated by the theoretical model, we make two assumptions. First, we assume that the conditional employment probability and wage in a market can be written as function of the average tax liability in that market only. Second, we assume that tax liabilities vary across individuals within a labor market according to some observable n. In our empirical application, this will correspond to the number of children in the household. The function describing participation decisions for individual m in labor market (e, s, t) can thus be written as:

$$\mathcal{K}_{m,e,s,t,n}(\mathbf{t}) = \tilde{\mathcal{K}}_{m,e,s,t,n}(p_{e,s,t}(T_{e,s,t}), w_{e,s,t}(T_{e,s,t}), T_{e,s,t,n}) \tag{17}$$

Taking a linear approximation to Equation (17) and adding state and year fixed effects, control variables and an error term, we get the following econometric specification:

$$k_{m,e,s,t,n} = \delta_{e,s} + \delta_{e,t} + T_{e,s,t,n}\beta_e + T_{e,s,t}\gamma_e + X_{m,e,s,t,n}\lambda_e + \nu_{m,e,s,t,n} \tag{18}$$

This equation implies that: $\beta_e = \frac{\partial \tilde{\mathcal{K}}_{m,e,s,t,n}}{\partial T_{e,s,t,n}}\Big|^{\text{Micro}}$, the micro participation effect, and $\gamma_e = \frac{\partial \tilde{\mathcal{K}}_{m,e,s,t,n}}{\partial p_{e,s,t}} \frac{\partial p_{e,s,t}}{\partial T_{e,s,t}} + \frac{\partial \tilde{\mathcal{K}}_{m,e,s,t,n}}{\partial w_{e,s,t}} \frac{\partial w_{e,s,t}}{\partial T_{e,s,t}}$. The macro effect is defined as the change in individual participation probabilities if the tax liabilities for all individuals in a labor market increase by one dollar. This simply corresponds to the response to a one dollar increase in $T_{e,s,t,n}$ as well as a one dollar increase in $T_{e,s,t}$ and therefore the macro effect can be obtained as: $\frac{d\tilde{\mathcal{K}}_{m,e,s,t,n}}{dT_{e,s,t}} = \beta_e + \gamma_e$.[33]

The market-level employment rate in market (e, s, t) is given by $\mathcal{H}_{e,s,t}(T_{e,s,t}) = p_{e,s,t}(T_{e,s,t}) \times \mathcal{K}_{e,s,t}(T_{e,s,t})$. Thus, the macro employment response is given by $\frac{\partial \mathcal{H}_{e,s,t}}{\partial T_{e,s,t}} = p_{e,s,t} \times \frac{\partial \mathcal{K}_{e,s,t}}{\partial T_{e,s,t}} + \mathcal{K}_{e,s,t} \times \frac{\partial p_{e,s,t}}{\partial T_{e,s,t}}$.

[33] Note that without income effects, $\frac{\partial \mathcal{K}_i}{\partial T_i} = \frac{\partial \mathcal{K}_i}{\partial b}$.In this case, only the difference in taxes and transfers between working and not working matters $T_i - T_i(0) = T_i + b$, and therefore $\frac{\partial \mathcal{K}_i}{\partial T_i} = \frac{\partial \mathcal{K}_i}{\partial b} = \frac{\partial \mathcal{K}_i}{\partial (T_i + b)}$. For our main specification, we will assume no income effects and therefore estimate directly $\frac{\partial \mathcal{K}_i}{\partial (T_i + b)}$ thereby using both variation in T_i and b to estimate the parameter of interest with maximum power. We tested whether the condition $\frac{\partial \mathcal{K}_i}{\partial T_i} = \frac{\partial \mathcal{K}_i}{\partial b}$ holds and found that the difference was very small and statistically insignificant. We therefore only report results under the no income effect assumption.

We will rely on a linear approximation for the market-level employment rate similar to Equation (18) and we will estimate the macro employment response in a way that is analogous to how we estimate the macro participation response.

Identification

To identify the parameters β_e and γ_e, two independent sources of exogenous variation in tax liabilities are needed. For the micro response β_e, we require variation in tax liabilities across individuals *within* the same labor market as can be seen from inspection of Equation (18). For the macro response $\beta_e + \gamma_e$, we require variation in average tax liabilities *between* labor markets.

The policy variation in the micro tax liability is displayed in Figure 1. This figures plots the 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. 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 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 colinear 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 2.

III.3 Empirical Results

III.3.a Micro-Macro Participation and Macro Employment Responses

The first panel of Table 2 shows the instrumental variables (IV) estimates for the micro participation responses to taxes and transfers. These correspond to estimates from a micro-level regression of participation on the tax liabilities controlling for number of children and labor market (education x state x year x month) fixed effects, age, age-squared, race, and ethnicity. The first column contains results for the full sample and each subsequent column contains a separate education group, starting with the lowest education group. Note that in interpreting these results that the tax liabilities are in units of \$1000. For all groups, we find a clear negative and statistically significant participation effects of taxes, consistent with the prior literature. For the full sample, we find that a \$1000 increase in taxes leads to a 3.4 percentage point reduction in participation which translates to an elasticity of 0.72.[34] We also see that the participation responses tend to be greater overall for women with lower levels of education. For example, the participation rates of single women who have less than a high school degree decrease by about 5.5 percentage points for every $1000 increase in tax liability which translates to an elasticity of 0.89. On the other hand, for bachelor degree holders, the corresponding elasticity is 0.46.

Our elasticity estimates, particularly for low educated single women, are somewhat large but they are within the range of elasticities that is reported in the literature.[35] 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.[36]

The macro participation responses are displayed in the second panel of Table 2. These correspond to empirical estimates from a macro-level (state-year cells) regression of participation rates on market-level tax liabilities, controlling for state and year fixed effects and percent black, percent hispanic, average age, average age-squared, average number of children and division-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 responses are smaller than the micro ones. For the individual education groups, the estimates are very noisy and not distinguishable from zero (or from the micro responses), but for the pooled estimates in column (1), the point estimate is statistically significant and about two-thirds the size of the micro response. The last panel shows the micro and macro participation responses from a regression of participation on both micro and macro tax liabilities, controlling for state and year fixed effects and individual controls. This allows for a statistical test of whether the micro and the macro response are statistically significantly different from each other. Although the spillover term is never statistically significant, we see that the magnitudes of the micro and macro tax elasticities generally line up with those in the first two panels.

Table 3 repeats the same set of regression results as Table 2, but considers employment, instead of participation, as the outcome variable. Recall that the sufficient statistic for the optimal tax formula is the macro employment response. For the fulll sample of single women, we see that a

[34] Following the theory, we take the marginal effect and multiply it by the ratio of the income gain from employment over the participation rate. For example, if we take the marginal effect of -0.034 and multiply it by the ratio $17.02/0.80, we get an elasticity of -0.72. We report the absolute value of the elasticities in the paper for simplicity.

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

[36] Most previous papers pool high school graduates and dropouts and thus find lower elasticies than for the high school dropouts alone. Furthermore, as we argue here, these papers confound micro and macro effects to some extent which can also explain why the reported elasticities may be a bit smaller.

$1000 increase in tax liability leads to a 2.1 percentage point reduction in employment rates. This implies a tax elasticity of 0.44. Similar to the results in Table 2, we see that the magnitude of the macro estimates generally decline with the level of education.

Table 4 provides a series of robustness tests where we drop the division-specific time trends and include alternative controls for pre-trends. The first column reports our baseline estimates for comparison. We see that the macro responses are very robust to controlling for region-specific time trends, state-specific time trends, division-by-year fixed effects, region-by-year fixed effects, and no controls for pre-trends.

Finally, the online Appendix reports the OLS regression results. We see that the OLS participation responses are attentuated 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 and suggests that endogeneity is a concern with OLS.

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 response and, in particular, may be quite misleading for the group of high school dropouts - a group that is particularly affected by policies such as AFDC/TANF or the EITC. Although this is some of the first evidence on the gap between micro and macro elasticities for various skill groups, an important limitation is that our macro effects are quite noisy and only statistically significant in the pooled estimations. This is likely due to a lack of statistical power given the limited policy variation at the state level over time. This highlights the need for future work that considers more reliable policy variation to identify macro effects.

IV Simulating the Optimal Tax Schedule

In this section we illustrate how the 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 derived above. In line with the empirical section, we focus on the schedule for the 4 education groups and the no-cross effects model with its restricted set of sufficient statistics.

IV.1 Optimal tax schedule with unemployment

To simulate the optimal tax schedule, we solve the system of first order conditions derived in the empirical section for the tax levels at different income levels. The system contains $N+2$ unknowns, the $i = 0 \ldots N$ tax levels T_i as well as the lagrange multiplier λ, and $N+2$ equations, the first order conditions (22) and (10) and the government budget constraint (1). Since we focus on the no-cross effects model, the first order conditions for the tax levels simplify to equation (13).[37] We partition the income distribution into discrete bins, corresponding to the zero income level and the 4 education groups in our empirical analysis. In order to solve the system of equations we also have to parameterize $g_i(T_i)$ and $h_i(T_i)$. Following Saez (2002) we parameterize g_i using the functional form: $g_i = \frac{1}{\lambda(w_i^0 - T_i^0)^\nu}$, where ν is the parameter describing society's parameter for redistriubtion. For comparability with Saez (2002) we set $\nu = 1$. We use a first order Taylor approximation to describe h_i, which should provide a reasonable approximation as long as the optimum is close to the current policy:

$$h_i = h_i^0 + \frac{\partial \mathcal{H}_i}{\partial T_i}\left(T_i - T_i^0\right) + \frac{\partial \mathcal{H}_i}{\partial b}\left(b_i - b_i^0\right). \tag{19}$$

Figure 3 shows the optimal tax and transfer schedule for the lowest 3 education groups using the employment and participation response estimates from our empirical section. Since the precision of our empirical estimates was highest when we pool all 4 education groups and since the pooled estimates are quite robust even for the macro responses, we rely on those for simulating the optimal tax schedule. The blue line shows the optimal tax schedule implied by our no-cross effects welfare formula, which relies on the the micro-macro participation gap to correct for spillovers. The figure also shows the corresponding optimal tax schedule implied by the pure extensive margin optimal tax formula in Saez (2002). The Saez (2002) formula relies only on employment responses but does not specify whether these are micro or macro responses. For the green line we implement the Saez (2002) formula using our micro employment response estimates, while for the red line we use the macro estimates. Compared to using the Saez formula with micro employment responses, our formula implies a substantially higher lump sum transfer and lower marginal tax rates. The Saez formula calibrated with macro employment responses implies larger transfers at the bottom than when micro employment responses are used for calibration. However both calibrations yield almost the same slope in the post-tax income profile and are quite a bit steeper than when spillovers are taken into account via our formula.

In Figure 4 we turn to some comparative statics to show how, holding the employment response constant, the macro-micro participation ratio affects the optimal tax schedule. The blue line shows the benchmark tax schedule from Figure 3 using our optimal tax formula with our main empirical estimates. The red line shows the optimal tax schedule using our formula when we double the macro-micro participation ratio but holding everything else constant. This captures a situation where the spillovers from an increase in employment taxes are positive (more labor market participants make it easier for people to find jobs). This makes the tax profile steeper

[37] In order to express the FOC for the benefit level in terms of sufficient statistics, we make two assumptions: a) benefits do not affect wages or job finding probabilities in any labor market and b) the social welfare function is linear in expected utilities (Benthamite Utilitarian). This can be viewed as an approximation that in practice likely does not make a big difference for the results.

and the optimal tax looks closer to an EITC like schedule. The green line on the other hand shows the optimal tax schedule when we cut the macro-micro participation ratio to 0.1, thus leading to large negative spillovers where the macro response is smaller than the micro response. This makes the overall tax profile flatter and the benefits to the non-employed larger.

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 and this has typically been explained by the possibility of job rationing at least in the short run. Figure 5 shows what happens to the optimal tax schedule if we double the macro participation and employment response. In panel a) we contrast the effect of doubling the macro responses using our optimal tax formula, while panel b) shows the effects using the Saez 2002 formula. While even in the Saez ()2002) formula a doubling of the employment response leads to significantly higher transfers at the bottom and a somewhat flatter tax schedule, this is much more pronounced using our formula that explicitly allows for unemployment and wage responses.

V 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 with fixed wages 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 redistribution 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. (2013). Second, it would be very interesting to study business cyle effects of taxation. For instance, the 2002 Farm Bill in the U.S. expanded the SNAP program at a time when unemployment rates were increasing. Whether such increases in the generosity of NIT-structured programs during recessions improves social welfare remains an open question. 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

Akerlof, G. A. and J. L. Yellen (1990). The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics 105*(2), 255–83.

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

Barro, R. J. and H. I. Grossman (1971). A General Disequilibrium Model of Income and Employment. *American Economic Review 61*(1), 82–93.

Blundell, R., J. C. Ham, and C. Meghir (1987). *Unemployment and female labour supply*. Number 149. Centre for Economic Policy Research.

Boadway, R. and J.-F. Tremblay (2013). Optimal income taxation and the labour market: An overview. *CESifo Economic Studies 59*(1), 93–148.

Brunello, G. and D. Sonedda (2007). Progressive taxation and wage setting when unions strategically interact. *Oxford Economic Papers 59*(1), 127–140.

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

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

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

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics 126*(2), 749–804.

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

Chetty, R., A. Guren, D. Manoli, and A. Weber (2012). Does indivisible labor explain the difference between micro and macro elasticities? a meta-analysis of extensive margin elasticities. *NBER Macroeconomics Annual 27*(1), 1–56.

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

Choné, P. and G. Laroque (2011). Optimal taxation in the extensive model. *Journal of Economic Theory 146*(2), 425–453.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *The Quarterly Journal of Economics 128*(2), 531–580.

Currie, J. and J. Gruber (1996). Health insurance eligibility and child health: lessons from recent expansions of the medicaid program. *Quarterly Journal of Economics 111*(2), 431–466.

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

Diamond, P. (1982). Wage Determination and Efficiency in Search Equilibrium. *Review of Economic Studies 49*(2), 217–27.

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

Eissa, N. and H. W. Hoynes (2006). Behavioral responses to taxes: Lessons from the eitc and labor supply. *Tax Policy and the Economy 20*, pp. 73–110.

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

Eissa, N. and J. B. Liebman (1996). Labor supply response to the earned income tax credit. *The Quarterly Journal of Economics 111*(2), pp. 605–637.

Farber, H. S. and R. G. Valletta (2013). Do extended unemployment benefits lengthen unemployment spells? evidence from recent cycles in the us labor market. Technical report, National Bureau of Economic Research.

Gelber, A. and J. Mitchell (2012). Taxes and time allocation: Evidence from single women and men. *Review of Economic Studies 79*, 863–897.

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

Hagedorn, M., F. Karahan, I. Manovskii, and K. Mitman (2013). Unemployment benefits and unemployment in the great recession: the role of macro effects. Technical report, National Bureau of Economic Research.

Ham, J. C. and K. T. Reilly (2002). Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data. *The American Economic Review 92*(4), 905–927.

Hansen, C. T., L. H. Pedersen, and T. Sløk (2000). Ambiguous effects of tax progressivity – theory and Danish evidence. *Labour Economics 7*(3), 335–347.

Herbst, C. M. (2008). Do social policy reforms have different impacts on employment and welfare use as economic conditions change? *Journal of Policy Analysis and Management 27*(4), 867–894.

Hersoug, T. (1984). Union wage responses to tax changes. *Oxford Economic Papers 36*(1), 37–51.

Holmlund, B. and A. S. Kolm (1995). Progressive taxation, wage setting and unemployment: theory and swedish evidence. *Swedish Economic Policy Review 2*, 423–460.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies 57*(2), 279–98.

Hoynes, H. W. (2000). Local Labor Markets And Welfare Spells: Do Demand Conditions Matter? *The Review of Economics and Statistics 82*(3), 351–368.

Hungerbühler, M., E. Lehmann, A. Parmentier, and B. Van der Linden (2006). Optimal redistributive taxation in a search equilibrium model. *Review of Economic Studies 73*(3), 743–767.

Jacquet, L., E. Lehmann, and B. Van der Linden (2014). Optimal income taxation with kalai wage bargaining and endogenous participation. *Social Choice and Welfare 42*(2), 381–402.

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

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. New York: Macmillan.

Kroft, K. (2008). Takeup, social multipliers and optimal social insurance. *Journal of Public Economics 92*(3-4), 623–668.

Kroft, K. and M. Notowidigdo (2014). Should unemployment insurance vary with the unemployment rate? theory and evidence. Technical report, National Bureau of Economic Research.

Lalive, R., C. Landais, and J. Zweimüller (2013). Market externalities of large unemployment insurance extension programs. Technical report, IZA Discussion Paper.

Landais, C., P. Michaillat, and E. Saez (2015). Optimal unemployment insurance over the business cycle. Working Paper 16526, National Bureau of Economic Research.

Lee, D. and E. Saez (2008). Optimal minimum wage policy in competitive labor markets. *NBER Working Papers 14 320*.

Lee, D. and E. Saez (2012). Optimal minimum wage policy in competitive labor markets. *Journal of Public Economics 96*(9 - 10), 739 – 749.

Lehmann, E., C. Lucifora, S. Moriconi, and B. Van der Linden (2014). Beyond the Labour Income Tax Wedge: The Unemployment-Reducing Effect of Tax Progressivity. IZA Discussion Papers 8276, Institute for the Study of Labor (IZA).

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

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

Malcomson, J. M. and N. Sartor (1987). Tax push inflation in a unionized labour market. *European Economic Review 31*(8), 1581 – 1596.

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

Marinescu, I. (2014). The general equilibrium impacts of unemployment insurance: Evidence from a large online job board.

McDonald, I. M. and R. M. Solow (1981). Wage Bargaining and Employment. *American Economic Review 71*(5), 896–908.

Meyer, B. and D. Sullivan (2004). The effects of welfare and tax reforms: The material well-being of single mothers in the 1980s and 1990s. *Journal of Public Economics 88*, 1387–1420.

Meyer, B. and D. Sullivan (2008). Changes in the consumption income and well-being of single mothers headed families. *American Economic Review 98*, 2221–2241.

Meyer, B. D. and D. T. Rosenbaum (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The Quarterly Journal of Economics 116*(3), pp. 1063–1114.

Michaillat, P. (2012). Do matching frictions explain unemployment? not in bad times. *The American Economic Review 102*(4), 1721–1750.

Michaillat, P. and E. Saez (2015). Aggregate Demand, Idle Time, and Unemployment. *The Quarterly Journal of Economics 130*(2), 507–569.

Mirrlees, J. (1999). Optimal marginal tax rates at low income. *mimeo*.

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

Mogstad, M. and C. Pronzato (2012). Are lone mothers responsive to policy changes? evidence from a workfare reform in a generous welfare state*. *The Scandinavian Journal of Economics 114*(4), 1129–1159.

Mortensen, D. T. and C. A. Pissarides (1999). New developments in models of search in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 39, pp. 2567–2627. Elsevier.

Naito, H. (1999). Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics 71*, 165–188.

Piketty, T., E. Saez, and S. Stantcheva (2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy 6*(1), 230–71.

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

Pissarides, C. A. (1985). Taxes, subsidies and equilibrium unemployment. *Review of Economic Studies 52*(1), 121–134.

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

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory, 2nd Edition*, Volume 1 of *MIT Press Books*. The MIT Press.

Pissarides, C. A. and B. Petrongolo (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature 39*(2), 390–431.

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

Rothstein, J. (2010). Is the eitc as good as an nit? conditional cash transfers and tax incidence. *American Economic Journal: Economic Policy 2*(1), pp. 177–208.

Rothstein, J. (2011). Unemployment insurance and job search in the great recession. Technical report, National Bureau of Economic Research.

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

Saez, E. (2002). Optimal income transfer programs: Intensive versus extensive labor supply responses. *The Quarterly Journal of Economics 117*(3), pp. 1039–1073.

Saez, E. (2004). Direct or indirect tax instruments for redistribution: short-run versus long-run. *Journal of Public Economics 88*, 503–518.

Schmieder, J., T. Von Wachter, and S. Bender (2012). The effects of extended unemployment insurance over the business cycle: Evidence from regression discontinuity estimates over 20 years. *The Quarterly Journal of Economics 127*(2), 701–752.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review 74*(3), 433–44.

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

A Theoretical Appendix

The Lagrangian associated to the government's program writes:

$$\Lambda(\mathbf{t}) \stackrel{\text{def}}{\equiv} \sum_{i=1}^{I} (T_i + b)\mathcal{H}_i(\mathbf{t}) - b - E + \frac{1}{\lambda}\Omega\left(\mathscr{U}_1(\mathbf{t}), ..., \mathscr{U}_I(\mathbf{t}), u(b)\right) \tag{20}$$

A.1 Derivation of Equation (10)

Differentiating (20) with respect to b gives:

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

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

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

Using $h_0 = 1 - \sum_{i=1}^{I} h_i$ and Equations (8) and (11) leads to (10). From $\frac{\partial \mathscr{C}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial T_i} - 1$ and for $j \neq i$, $\frac{\partial \mathscr{C}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial T_i}$, the sum of (9) for all T_j minus Equation (10) leads to:

$$\begin{aligned} 0 \quad = \quad & \sum_{i=1}^{0} h_i + \sum_{i=1}^{I} (T_i + b)\left(\sum_{j=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_j} - \frac{\partial \mathcal{H}_i}{\partial b}\right) - \left(g_0 h_0 + \sum_{i=1}^{I} g_i h_i\right) \\ + \quad & \sum_{i=1}^{I} g_i h_i \left(\sum_{j=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} - \frac{\partial \mathscr{W}_i}{\partial b}\right) + \sum_{i=1}^{I} g_i h_i \frac{u(c_i) - d_i - u(b)}{u'(c_i)}\left(\sum_{j=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_j} - \frac{\partial \mathscr{P}_i}{\partial b}\right) \end{aligned} \tag{21}$$

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 $\sum_{i=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial b}$, $\sum_{i=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_i} = \frac{\partial \mathscr{P}_i}{\partial b}$ and $\sum_{i=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_i} = \frac{\partial \mathcal{H}_i}{\partial b}$. Plugging these equalities in (21) leads to: $g_0 h_0 + \sum_{i=1}^{I} g_i h_i = 1$.

A.2 Derivation of Equation (14)

Let $\mathcal{A}$ denotes the square matrix of rank I whose term in row j and column i is $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j}\ \frac{u(c_i) - d_i - u(b)}{p_i\ u'(c_i)}$. The optimal tax formula (9) can be rewritten in matrix notations:

$$\mathbf{0} = \underbrace{\mathbf{h}}_{\text{Mechanical effect}} + \underbrace{\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b})}_{\text{Behavioral effects}} + \underbrace{\mathcal{A} \cdot (\mathbf{g}\ \mathbf{h})}_{\text{Social Welfare effects}} \tag{22}$$

However, Equation (5) implies that: $\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} = -\mathcal{A} \cdot \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\Big|^{\text{Micro}}$. Moreover, from $\mathcal{K}_i(\mathbf{t}) = \hat{\mathcal{K}}_i(\mathscr{U}(\mathbf{t}))$, we get that: $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} = \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$ and $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\Big|^{\text{Micro}} = \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\Big|^{\text{Micro}} \cdot \frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$. We thus get that:

$$-\mathcal{A} = \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \left(\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\Big|^{\text{Micro}}\right)^{-1} = \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\Big|^{\text{Micro}}\right)^{-1}$$

whenever $\frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$ is invertible, in which case Equation (22) can be rewritten as (14).

A.3 The Matching model

We consider a matching economy where on each labor market i, the constant returns to scale matching function gives the employment level h_i as a function $\mathcal{M}_i(\nu_i, k_i)$ of the number ν_i of vacancies posted and the number k_i of participating job seekers (Pissarides and Petrongolo, 2001). Creating a jobs costs $\kappa_i > 0$ and generates output $y_i > \kappa_i$ when a worker is recruited. Hence, the different types of labor are perfect substitutes.

Each vacancy is matched with probability $q_i = Q_i(\theta_i) \overset{\text{def}}{\equiv} \frac{\mathcal{M}_i(\nu_i,k_i)}{\nu_i} = \mathcal{M}_i(1, 1/\theta_i)$, which is decreasing in tightness $\theta_i \overset{\text{def}}{\equiv} \nu_i/k_i$. Firms create jobs whenever the expected profit $q_i(y_i - w_i) - \kappa_i$ is positive. As more vacancies are created, tightness decreases until the free entry condition $q_i(y_i - w_i) = \kappa_i$ is verified. The conditional employment probability is an increasing function of tightness through $p_i = P(\theta_i) \overset{\text{def}}{\equiv} \frac{\mathcal{M}_i(\nu_i,k_i)}{k_i} = \mathcal{M}_i(\theta_i, 1)$. Therefore, the conditional probability p_i is a decreasing function of the gross wage through $p_i = P_i\left(Q_i^{-1}\left(\frac{\kappa_i}{y_i - w_i}\right)\right)$, which determines the labor demand function $p_i = \mathscr{L}_i(w_i)$.

Under risk neutrality and proportional bargaining (15), one has for any $j \neq i$ that $\frac{\partial \mathscr{W}_i}{\partial T_j} = 0$, thereby $\frac{\partial \mathscr{P}_i}{\partial T_j} = 0$ from $p_i = \mathscr{L}_i(w_i)$, and finally $\frac{\partial \mathscr{U}_i}{\partial T_j} = 0$ from (5). Moreover, we get from $p_i = \mathscr{L}_i(w_i)$ and (5) that:

$$\frac{\partial \mathscr{U}_i}{\partial T_i} = \left[-1 + \frac{\partial \mathscr{W}_i}{\partial T_i}\left(1 + \frac{w_i}{p_i}\frac{\partial \mathscr{P}_i}{\partial w_i}\frac{w_i - T_i - d_i - b}{w_i}\right)\right] p_i$$

As $\mu_i \in (0,1)$ denote the elasticity of the matching function with respect to the number of job-seekers, we get $\frac{dp_i}{p_i} = (1-\mu_i)\frac{d\theta_i}{\theta_i}$ and $\frac{dq_i}{q_i} = -\mu_i \frac{d\theta_i}{\theta_i}$, so $\frac{dp_i}{p_i} = -\frac{1-\mu_i}{\mu_i}\frac{dq_i}{q_i}$. Log-differentiating the free-entry condition $k_i = q_i\ (y_i - w_i)$ leads to $\frac{dq_i}{q_i} = \frac{w_i}{y_i - w_i}\frac{dw_i}{w_i}$. So, we get $\frac{dp_i}{p_i} = -\frac{1-\mu_i}{\mu_i}\frac{w_i}{y_i - w_i}\frac{dw_i}{w_i}$, i.e: $\frac{w_i}{p_i}\frac{\partial \mathscr{P}_i}{\partial w_i} = -\frac{1-\mu_i}{\mu_i}\frac{w_i}{y_i - w_i}$ and:

$$\frac{\partial \mathscr{U}_i}{\partial T_i} = \left[-1 + \frac{\partial \mathscr{W}_i}{\partial T_i}\left(1 - \frac{1-\mu_i}{\mu_i}\frac{w_i - T_i - d_i - b}{y_i - w_i}\right)\right] p_i$$

Equation (15) implying that $\frac{w_i - T_i - d_i - b}{y_i - w_i} = \frac{\beta_i}{1-\beta_i}$ and $\frac{\partial \mathscr{W}_i}{\partial T_i} = 1 - \beta_i$, we get:

$$\frac{\partial \mathscr{U}_i}{\partial T_i} = \left[-1 + (1-\beta_i)\left(1 - \frac{1-\mu_i}{\mu_i}\frac{\beta_i}{1-\beta_i}\right)\right] p_i = \frac{\beta_i}{\mu_i}\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \tag{23}$$

when $\mu_i > 0$ and $\beta_i < 1$, which ends the proof of Proposition 3.

A.4 The case without unemployment responses

In this Appendix, we consider the case where the conditional employment probability is exogenous at $p_i \in (0,1]$ (so $\frac{\mathbf{d}\mathscr{P}}{\mathbf{dT}} = \mathbf{0}$) and where the different types of labor are substitutable. More specifically, we assume that the different types of labor h_i and capital Z produce a numeraire good sold in a perfectly competitive product market under a constant returns to scale technology $F(h_1, ..., h_I, Z)$.[38] We furthermore assume the rate of return to capital, $r > 0$, is exogenous. The latter assumption can be viewed either by considering a small open economy and assuming perfect capital mobility, or by considering the steady state of a closed economy with infinite horizon

[38] We hence generalize Saez (2002) who considered perfect substitution across the difference types of labor through the production function: $F(h_1, ..., h_I) = \sum_{i=1}^{I} w_i\, h_i$, where w_i stands both for the productivity of labor in occupation i and for the wage in the corresponding labor market.

savers. The assumptions of exogenous unemployment rates and constant returns to scale seem plausible in the long run, even though they ruled out job rationing considered by Landais et al. (2015) which are plausible in the short run. We then get that:

Proposition 4. *If the unemployment rates are exogenous, the production function exhibits constant returns to scale and $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ is invertible, the optimal tax schedule is given by:*

$$0 = (1-g_j)h_j + \sum_{i=1}^{I}(T_i+b)\ \frac{\partial \mathcal{H}_i}{\partial T_j}\bigg|^{Micro} \tag{24}$$

and depends only on microeconomic employment responses.

Proof: In the absence of unemployment responses to taxation $\frac{\partial \mathscr{P}_i}{\partial T_j} = 0$, the matrix $\mathcal{A}$ of corrective terms $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j}\frac{u(c_i)-d_i-u(b)}{p_i\, u'(c_i)}$ coincides with $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$. We thus get: $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} = -\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\Big|^{\text{Micro}}$ and $\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} = -\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\Big|^{\text{Micro}}$. Equation (14) then successively leads to:

$$\begin{aligned}
\mathbf{0} &= \mathbf{h} - \frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\bigg|^{\text{Micro}} \cdot (\mathbf{T}+\mathbf{b}) + \frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\bigg|^{\text{Micro}} \cdot \left(\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\bigg|^{\text{Micro}}\right)^{\mathbf{-1}} \cdot (\mathbf{g}\,\mathbf{h}) \\
\mathbf{0} &= \mathbf{h} - \frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\bigg|^{\text{Micro}} \cdot (\mathbf{T}+\mathbf{b}) + \frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot (\mathbf{g}\,\mathbf{h}) \\
\mathbf{0} &= \left(\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}\right)^{\mathbf{-1}} \cdot \mathbf{h} - \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\bigg|^{\text{Micro}} \cdot (\mathbf{T}+\mathbf{b}) + \mathbf{g}\,\mathbf{h}
\end{aligned} \tag{25}$$

where the last equality requires the matrix $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ to be invertible.

Moreover, the firm's profit function verifies $\Pi(w_1,...,w_I,r) \overset{\text{def}}{\equiv} \max_{h_1,...,h_I,Z} \quad F(h_1,...,h_I,Z) - \sum_{i=1}^{I} w_i\, h_i - r\,Z$. Applying the envelope theorem leads to $\frac{\partial \Pi}{\partial w_i} = -h_i$, thereby $d\Pi = -\sum_{i=1}^{I} h_i\, dw_i - Z\, dR$. Because of perfect competition and constant returns to scale, we get that $d\Pi = 0$, which together with the assumption of an inelastic return of capital leads to $0 = \sum_{i=1}^{I} h_i \frac{\partial \mathscr{W}_i}{\partial T_j}$. In matrix notation, this implies that $\mathbf{h}$ is an eigenvector of Matrix $\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}$ associated to eigenvalue 0. Hence, $\mathbf{h}$ is an eigenvector of Matrix $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ associated to eigenvalue -1, so $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}} \cdot \mathbf{h} = -\mathbf{h}$ and eventually $\left(\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}\right)^{\mathbf{-1}} \cdot \mathbf{h} = -\mathbf{h}$. Therefore Equation (25) simplifies to:

$$\mathbf{0} = \mathbf{1} - \mathbf{g}\,\mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\bigg|^{\text{Micro}} \cdot (\mathbf{T}+\mathbf{b})$$

which corresponds to (24).

This result may look surprising and is also due to the specific representation of the labor supply responses along the intensive margin in the occupation model of Saez (2002). Stiglitz (1982), Naito (1999) propose alternatively a two-skills version of the Mirrlees model with intensive labor supply responses where low skilled and high skilled labor are imperfect substitutes. Stiglitz (1982) shows that the labor supply of the high skilled workers needs to be upward distorted (negative marginal tax rate for high skilled workers), unless the elasticity of substitution across the two types of labor is infinite. This result of Stiglitz (1982) looks at odds with the result above. Saez (2004) explains this discrepancy by the fact that in Stiglitz (1982) when a high skill worker earns the gross income intended to a low-skilled one, he does so keeping her high skill productivity. In other words, a worker's skill is portable across the different income levels in Stiglitz (1982) but not in Saez (2004). Therefore, a change in the low skilled gross wage affects the self-selection incentive constraint in Stiglitz (1982) and Naito (1999), as well as in the continuous income model of Rothschild and Scheuer (2013), while in the occupation model of Saez (2004) and Lee and Saez (2008), when an individual works in a low-skilled job, she has a low productivity. The occupation model captures not only extensive (participation) responses but also educational choice along the intensive margin in the long-run while the models of Stiglitz (1982) and Naito (1999) focus on the short-run hours of work and in-work effort responses along the intensive margin. $\square$

Table 1: Variable Means for Single Women

	(1) Full Sample	(2) High School Dropout	(3) High School Graduate	(4) Some College	(5) Bachelors Degree Plus
Panel A: Demographics					
Age	34.5	33.6	33.9	34.5	35.9
No Children Percent	69.1	59.6	65.8	67.0	82.3
1 Child Percent	16.1	16.9	17.8	18.0	10.7
2 Children Percent	9.6	12.3	10.6	10.3	5.5
3+ Children Percent	5.2	11.2	5.8	4.7	1.6
Mean Years of Education	13.1	9.3	12	13.3	16.6
Percent Black	19.0	24.6	21.5	18.7	12.2
Percent Hispanic	12.6	29.9	12.2	10.0	6.1
Panel B: Labor Force Status					
Labor Force Participation Rate (k_i)	80.4	55.2	78.3	85.3	92.1
Employment Rate (h_i)	74.6	45.9	71.4	80.2	89.2
Unemployment Rate ($1 - p_i$)	7.9	17.1	8.9	6.1	3.2
Panel C: Income, Taxes and Transfers (Real 2010 Dollars)					
Imputed Pre-tax Wage Earnings	21897	10022	16925	21503	36547
Net Taxes: No Children	5731	1666	3717	5092	10438
Net Taxes: 2 Children	505	-1991	-1286	375	9145
AFDC/TANF and Food Stamps: No Children	468	1356	638	359	8
AFDC/TANF and Food Stamps: 2 Children	3266	7179	3666	1944	95
Net Tax and Transfers (T_i): No Children	5262	311	3079	4733	10430
Net Tax and Transfers (T_i): 2 Children	-2761	-9169	-4951	-1569	9050
Net Tax and Transfers (b): Zero Income, No Children	-2072	-2055	-2069	-2077	-2076
Net Tax and Transfers (b): Zero Income, 2 Children	-11489	-11550	-11442	-11480	-11558
AFDC/TANF Recipiency Rate for Mothers: Pre-1996	27	49	25	17	3
AFDC/TANF Recipiency Rate for Mothers: Post-1996	9	21	10	6	1
Food Stamp Recipiency Rate: Pre-1996	17	41	19	10	1
Food Stamp Recipiency Rate: Post-1996	17	41	23	15	2
Number of observations	1008322	139260	334360	300246	234456

Notes: The sample is restricted to single women aged 18-55. All dollar figures are in real 2010 dollars.
Data used in each column are restricted to women 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 recipiecy 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 Participation Responses to Changes in Taxes and Benefits
Instrumental Variable Regressions

	(1) Full Sample	(2) HS Drop	(3) HS Grad	(4) Some College	(5) Bachelor Degree Plus
Micro Response					
Taxes Plus Benefit $\frac{\partial \mathcal{K}_i}{\partial T_i}\Big\rvert_{\mathbf{w},\mathbf{p}}$	-0.034	-0.055	-0.031	-0.018	-0.016
	[0.002]***	[0.003]***	[0.002]***	[0.002]***	[0.003]***
Num. Obs	1007760	138766	334359	300242	234393
Mean of Dep. Var.	0.80	0.55	0.78	0.85	0.92
Inc Gain from Employment (2010USD)	17019.6	8884.6	13928.7	17111.0	26142.4
Tax Elasticity	-0.72	-0.89	-0.55	-0.37	-0.46
Macro Response					
Taxes Plus Benefit $\frac{\partial \mathcal{K}_i}{\partial T_i}$	-0.021	-0.042	-0.041	-0.020	-0.009
	[0.010]**	[0.031]	[0.028]	[0.021]	[0.006]
Num. Obs	5712	1428	1428	1428	1428
Mean of Dep. Var.	0.79	0.56	0.80	0.87	0.93
Inc Gain from Employment (2010USD)	15409.7	8202.6	13176.6	16060.0	24199.7
Tax Elasticity	-0.42	-0.61	-0.68	-0.37	-0.24
Simultaneous Estimation					
Taxes Plus Benefit $\frac{\partial \mathcal{K}_i}{\partial T_i}\Big\rvert_{\mathbf{w},\mathbf{p}}$	-0.034	-0.056	-0.031	-0.018	-0.015
	[0.002]***	[0.003]***	[0.002]***	[0.002]***	[0.003]***
Wedge Term $\frac{\partial \mathcal{K}_i}{\partial T_i} - \frac{\partial \mathcal{K}_i}{\partial T_i}\Big\rvert_{\mathbf{w},\mathbf{p}}$	0.006	0.025	-0.040	0.001	0.008
	[0.016]	[0.021]	[0.069]	[0.019]	[0.008]
Num. Obs	1008322	139260	334360	300246	234456
Tax Elasticity	-0.73	-0.91	-0.56	-0.37	-0.42
Macro Tax Elasticity	-0.61	-0.50	-1.27	-0.35	-0.20

Notes: Standard errors clustered on state level (** P<.05, *** P<.01)). 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. Data used in each column are restricted to women with the education level in the column header.Tax 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 first panel of 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 data used in the second panel are collapsed to the state-year cell observations, each cell recieves 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 division time trends. The third panel of regressions use individual level data and include controls for age, age-squared race, ethnicity and fixed effects for number of children and state and year and CPS division time trends. All controls in the first column of each panlel are interacted with education group.

Table 3: Micro and Macro Employment Responses to Changes in Taxes and Benefits
Instrumental Variable Regressions

	(1) Full Sample	(2) HS Drop	(3) HS Grad	(4) Some College	(5) Bachelor Degree Plus
Micro Response					
Taxes Plus Benefit $\left.\frac{\partial \mathcal{H}_i}{\partial T_i}\right\|_{\mathbf{w},\mathbf{p}}$	-0.032	-0.049	-0.031	-0.019	-0.018
	[0.002]***	[0.003]***	[0.002]***	[0.002]***	[0.004]***
Num. Obs	1007760	138766	334359	300242	234393
Mean of Dep. Var.	0.75	0.46	0.71	0.80	0.89
Inc Gain from Employment (2010USD)	17019.6	8884.6	13928.7	17111.0	26142.4
Tax Elasticity	-0.74	-0.94	-0.60	-0.41	-0.52
Macro Response					
Taxes Plus Benefit $\frac{\partial \mathcal{H}_i}{\partial T_i}$	-0.021	-0.044	-0.025	-0.027	-0.009
	[0.010]**	[0.032]	[0.031]	[0.026]	[0.007]
Num. Obs	5712	1428	1428	1428	1428
Mean of Dep. Var.	0.73	0.46	0.73	0.81	0.90
Inc Gain from Employment (2010USD)	15409.7	8202.6	13176.6	16060.0	24199.7
Tax Elasticity	-0.44	-0.77	-0.45	-0.54	-0.23
Simultaneous Estimation					
Taxes Plus Benefit $\left.\frac{\partial \mathcal{H}_i}{\partial T_i}\right\|_{\mathbf{w},\mathbf{p}}$	-0.033	-0.049	-0.031	-0.020	-0.017
	[0.002]***	[0.003]***	[0.002]***	[0.002]***	[0.004]***
Wedge Term $\frac{\partial \mathcal{H}_i}{\partial T_i} - \left.\frac{\partial \mathcal{H}_i}{\partial T_i}\right\|_{\mathbf{w},\mathbf{p}}$	0.006	0.023	-0.015	-0.001	0.005
	[0.018]	[0.024]	[0.057]	[0.023]	[0.011]
Num. Obs	1008322	139260	334360	300246	234456
Tax Elasticity	-0.75	-0.95	-0.61	-0.42	-0.49
Macro Tax Elasticity	-0.61	-0.50	-0.90	-0.44	-0.34

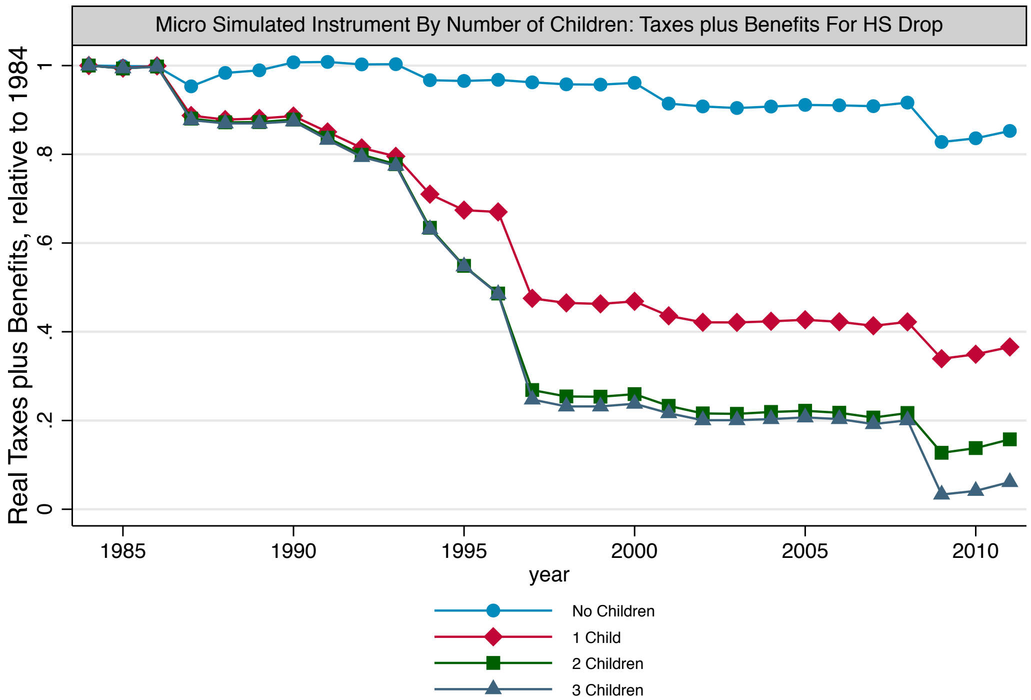
Notes: Standard errors clustered on state level (** P<.05, *** P<.01)). The sample is restricted to single women aged 18-55. The data include March CPS for 1984-2011 and Outgoing Rotations Groups for 1994-2010. Data used in each column are restricted to women with the education level in the column header.Tax 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 first panel of 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 data used in the second panel are collapsed to the state-year cell observations, each cell recieves 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 division time trends. The third panel of regressions use individual level data and include controls for age, age-squared race, ethnicity and fixed effects for number of children and state and year and CPS division time trends. All controls in the first column of each panlel are interacted with education group.

Table 4: Macro Participation and Employment Responses to Changes in Tax Liability
IV Full Sample Robustness Checks

	(1) Division Time Trend	(2) Region Time Trend	(3) State Time Trend	(4) Div X Year FE	(5) Reg X Year FE	(6) No Pre-Trends
Macro Participation Response						
Taxes Plus Benefit $\frac{\partial \mathcal{K}_i}{\partial T_i}$	-0.021	-0.022	-0.024	-0.021	-0.025	-0.029
	[0.010]**	[0.012]*	[0.013]*	[0.011]*	[0.014]*	[0.012]**
Num. Obs	5712	5712	5712	5712	5712	5712
Mean of Dep. Var.	0.79	0.79	0.79	0.79	0.79	0.79
Income Gain from Employment	15409.7	15409.7	15409.7	15409.7	15409.7	15409.7
Tax Elasticity	-0.42	-0.44	-0.47	-0.40	-0.49	-0.57
Macro Employment Response						
Taxes Plus Benefit $\frac{\partial \mathcal{H}_i}{\partial T_i}$	-0.021	-0.019	-0.032	-0.018	-0.019	-0.026
	[0.010]**	[0.011]*	[0.016]**	[0.011]	[0.013]	[0.012]**
Num. Obs	5712	5712	5712	5712	5712	5712
Mean of Dep. Var.	0.73	0.73	0.73	0.73	0.73	0.73
Income Gain from Employment	15409.7	15409.7	15409.7	15409.7	15409.7	15409.7
Tax Elasticity	-0.44	-0.41	-0.68	-0.37	-0.40	-0.56

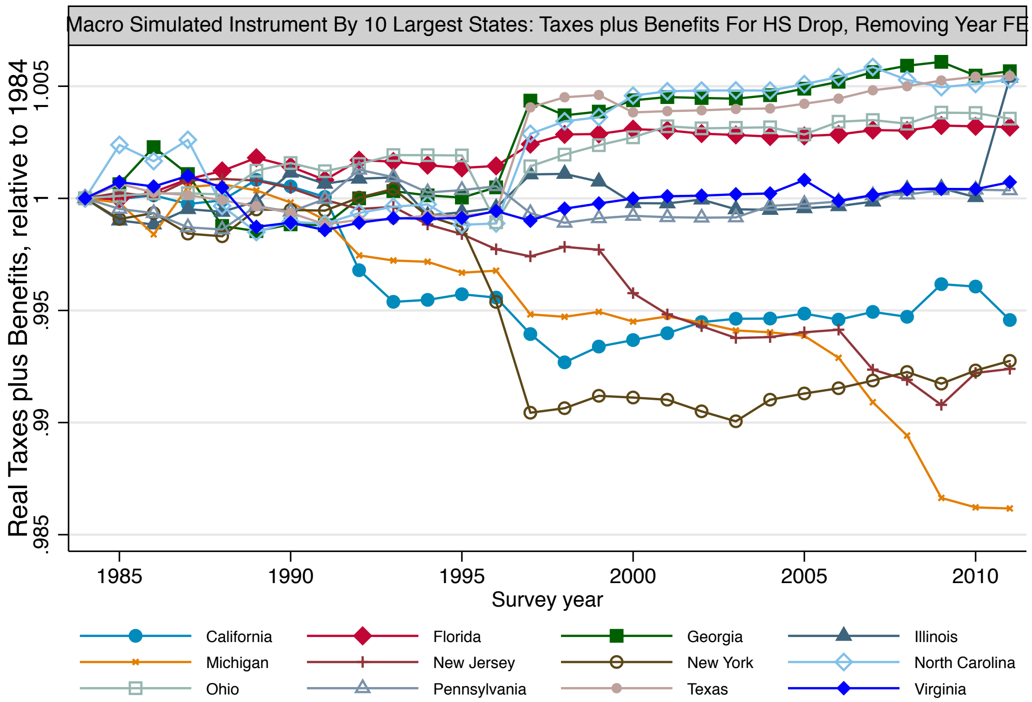
Notes: Standard errors clustered on state level (** P<.05, *** P<.01)). The sample is restricted to single women aged 18-55. The data include March CPS for 1984-2011 and Outgoing Rotations Groups for 1994-2010. The full sample is used for all columns, each column include controls described in the column header. Tax 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 data are collapsed to the state-year cell, each cell recieves 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. All controls are interacted with education group.

Figure 1: Micro Variation in Taxes plus Benefits



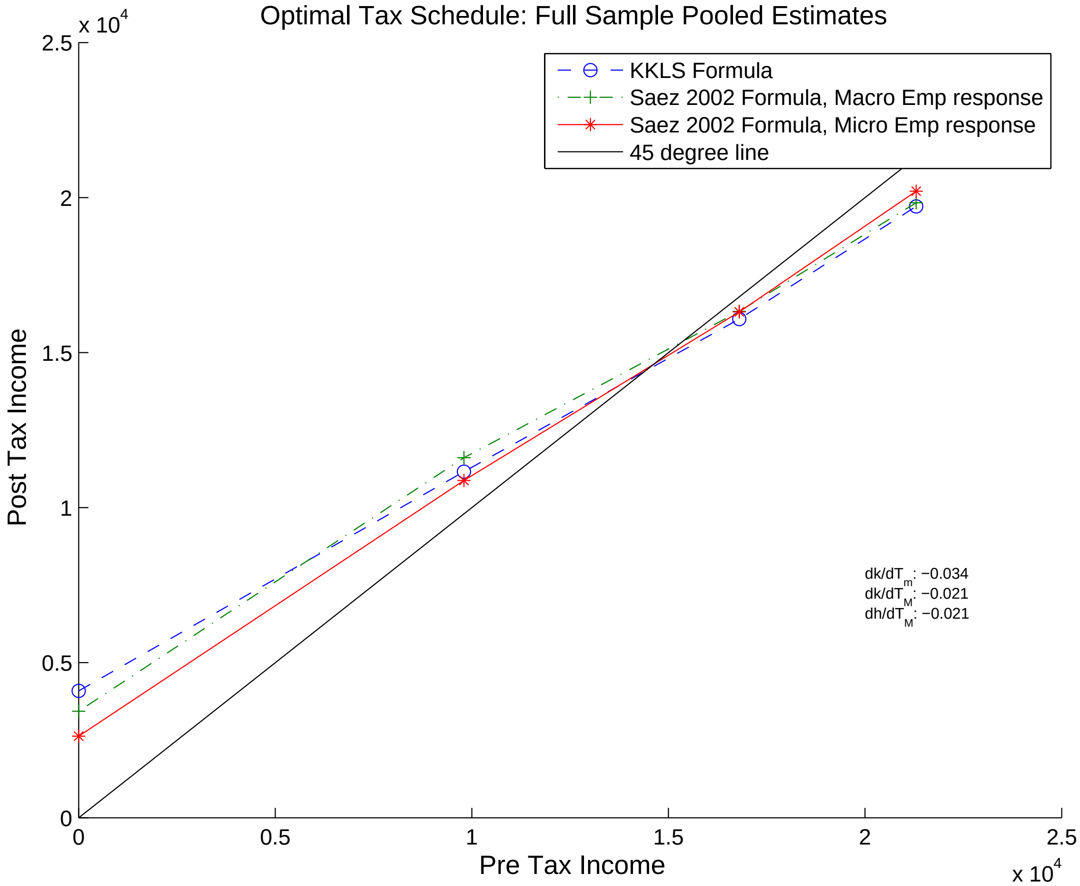
Notes: The figure shows the variation in taxes plus benefits for high school dropouts by number of children normalized such that 1984 equals one.

Figure 2: Micro Variation in Taxes plus Benefits



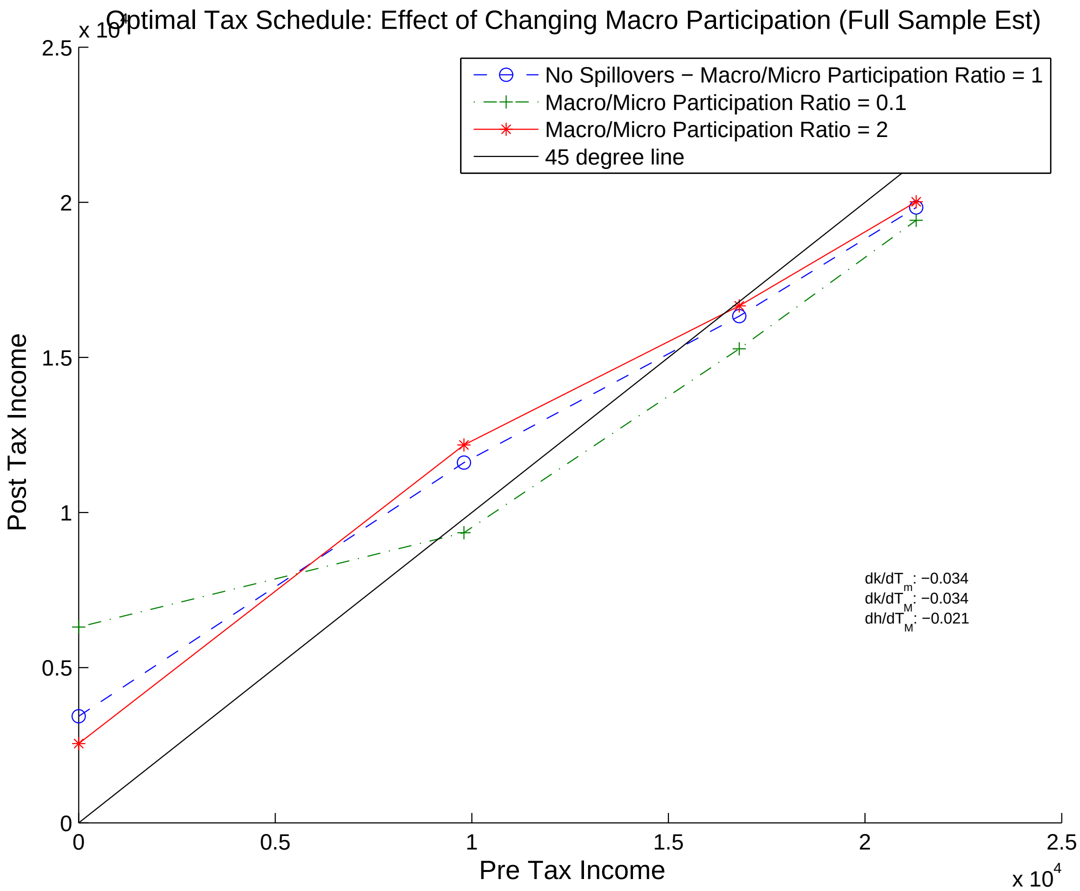
Notes: This figure shows residuals from a regression of year fixed effects on the macro level taxes plus benefits with state means added back to the residual, then normalized such that 1984 equals one.

Figure 3: Optimal Tax and Transfer Schedule Comparing KKLS Formula with Saez (2002) Formula



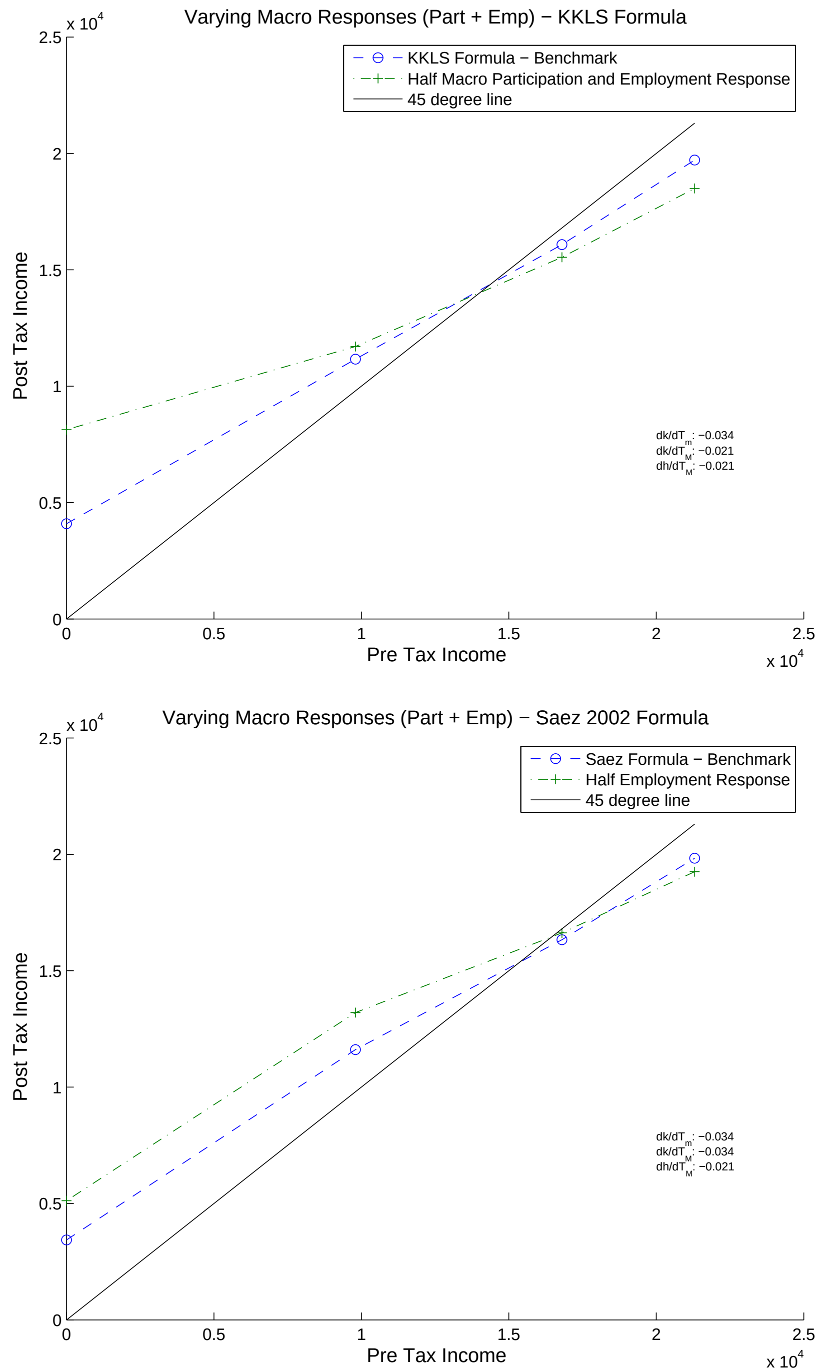
Notes: Simulations of the optimal tax and transfer schedule under alternate assumptions on employment and participation responses. Distribution of the 4 income groups is calibrated using CPS data and corresponds to the 4 education groups in the empirical section. The figure uses the participation and employment responses estimated in the paper. Since macro responses are not well identified for college graduate,they are set to the corresponding micro responses. The blue line uses the optimal welfare formula derived in this paper. The green line uses the Saez (2002) formula based on the estimated macro responses in this paper, while the red line uses the estimated micro employment responses in this paper.

Figure 4: Optimal Tax and Transfer Schedule: Comparative Statics with Varying Macro Participation Response



Notes: Simulations of the optimal tax and transfer schedule under alternate macro participation responses. Distribution of the 4 income groups is calibrated using CPS data and corresponds to the 4 education groups in the empirical section. The blue line corresponds to the benchmark simulation using the KKLS optimal tax formula and the estimated, participation and employment responses. The red line shows the optimal tax schedule when the macro participation response is doubled. The green line shows the optimal tax schedule when the macro participation response is halfed.

Figure 5: Optimal Tax and Transfer Schedule: Comparative Statics with Varying Macro Participation and Employment Responses



Notes: Simulations of the optimal tax and transfer schedule under alternate macro participation responses. Distribution of the 4 income groups is calibrated using CPS data and corresponds to the 4 education groups in the empirical section. The top figure uses the KKLS optimal tax formula, the bottom figure the Saez (2002) optimal tax formula. The blue line corresponds to the benchmark simulation using the estimated, participation and employment responses. The red line shows the tax schedule when the macro participation and employment responses are doubled. The green line shows the tax schedule when the macro participation and employment response is halfed.