CESifo WORKING PAPERS

6686
2017

October 2017

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

Kory Kroft, Kavan Kucko, Etienne Lehmann, Johannes Schmieder

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6686*
Category 1: Public Finance

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

Abstract

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

Kory Kroft
University of Toronto
150 St. George Street
Canada – 354 Toronto, ON
kory.kroft@utoronto.ca

Kavan Kucko
Boston University
Department of Economics
USA – Boston, MA 02215
kavan.kucko@gmail.com

Etienne Lehmann
CRED Université Panthéon-Assas Paris II
12 Place du Panthéon
France – 75 231 Paris Cedex
etienne.lehmann@gmail.com

Johannes Schmieder
Boston University
Department of Economics
USA – Boston, MA 02215
johannes@bu.edu

September 20, 2017
We would like to thank Felix Bierbrauer, Pierre Cahuc, David Card, Raj Chetty, Sebastian Findeisen, Peter Funk, Robert Gary-Bobo, Emmanuel Hansen, Bas Jacobs, Laurence Jacquet, Hilary Hoynes, Henrik Kleven, Bas Jacobs, Claus Thustrup Kreiner, Patrick Kline, Rafael Lalive, Kevin Lang, Thomas Le Barbanchon, Isabelle Mejean, Jean-Baptiste Michau, Austin Nichols, Matthew Notowidigdo, Claudia Olivetti, Daniele Paserman, Thomas Piketty, Jukka Pirttilä, Julien Prat, Jesse Rothstein, Dominik Sacks, Emmanuel Saez, Stefanie Stantcheva, Aleh Tsyvinski, Bruno Van der Linden, Owen Zidar as well as seminar participants at Boston University, McGill University, University of Connecticut, University of Toronto, CRED-University Paris II, CREST, TEPP, University of Cologne, Queens University, THEMA Cergy-Pontoise, SOLE Meeting 2013, the NTA Meeting 2014, Berkeley, Hamilton,Wilfried Laurier, CEPR, SED meeting 2015, NBER Summer Institute, University of Chile, Lille 1, Nancy, Lausanne, Mannheim, Ecole Polytechnique, AFSE 2016, PET 2016, IIPF 2016 and PSE for many helpful comments on this project. All errors are our own. Etienne Lehmann is also research fellow at IZA, CESifo, IDEP and CEPR and is junior member of IUF. A part of this research was realized while Etienne Lehmann was visiting the Center for Equitable Growth at UC Berkeley whose financial support is gratefully acknowledged.

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 and income (Eissa and Liebman, 1996, Meyer and Rosenbaum, 2001, Eissa and Hoynes, 2006, Gelber and Mitchell, 2012, Hoynes and Patel, 2015) 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 welfare weight on the working poor is greater than 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 U.S. 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. In fact, recent work by Bitler et al. (2014) shows that for single women, the EITC does not provide much protection during economic downturns. Furthermore, even in a full employment economy, the assumption of fixed wages may be implausible (Rothstein, 2010). It is also worth noting that these assumptions rule out any labor market spillover effects of government policies. Since anyone can find a job at all times, there is no mechanism by which a boost to the labor force could "crowd out" job finding. Thus, these assumptions are at odds with the growing body of evidence that suggest, especially during times when unemployment is high, government policies may induce substantial spillover effects, particularly at the bottom end of the income distribution. It is desirable to have a theoretical framework that can account for the presence of these spillovers.

The goal of this paper is to relax the fixed wage and full employment assumptions and reassess whether the optimal income tax features an EITC-like structure at the bottom, as in Saez (2002). The paper makes two key contributions, one theoretical and one empirical. Theoretically, we derive a sufficient statistics optimal tax formula in a general model that incorporates unemployment and wage responses to taxation. 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 Saez (2002) where all active individuals are effectively working. 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". Rather than specifying the full structure of the labor market, we pursue a sufficient statistics approach (Chetty, 2009) by allowing wages and the "conditional employment probability" - the fraction of participating individuals that are effectively working (i.e. one minus the unemployment rate) - to depend in a reduced-form way on taxes. 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 micro responses, macro responses allow wages and conditional employment probabilities in each labor market to respond to a change in taxes. When we consider a restricted version of the model, whereby tax liabilities in one market do not affect wages, conditional employment probabilities, and labor supply in other occupations (what we label the "no-cross effects" model), 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.[2] 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 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.[3] 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 utilities as well, these market spillovers are entirely captured by the ratio of macro over micro participation responses. This is related to results in Kroft

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

[2] The no-cross effects model resembles the pure extensive model in Saez (2002), but additionally allows for unemployment and wage responses to changes in tax liabilities in the same occupation.

[3] 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.

(2008) and Landais et al. (2016) who show that to evaluate optimal unemployment insurance (UI), it is important to estimate the ratio of the micro and macro take-up and duration elasticities in the presence of spillover effects, respectively. We view this as a very general result and one that extends beyond income taxation and unemployment insurance. Intuitively, any government policy that affects expected utility will affect labor force participation. Thus, participation responses can be thought of as a useful ”revealed preference” guide to measuring the welfare effects of government policies.

The optimal tax formulas structure our empirical strategy which estimates the sufficient statistics that are inputs to the optimal tax formula using a standard quasi-experimental research design. Most of the U.S. literature on labor supply responses to taxation and transfer has focused on single women, who are most likely to be at the margin of participating in the labor market and are thereby most affected by tax and transfer policies at the bottom of the income distribution, in particular the EITC. We largely follow this approach, but also include single men. Focusing on unmarried individuals avoids complications due to joint taxation.[4] 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 and year. This largely mirrors the definition of labor markets in Rothstein (2010). To identify the micro participation response, we rely on expansions to the federal EITC which differentially affected single individuals with and without children. For the macro participation and employment responses, 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 in spirit 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 individuals, is 0.57. This generally lines up with the range of estimates reported in the literature (Eissa et al., 2008). Our estimate of the macro participation and employment elasticity is 0.48 and 0.42, respectively. Finally, we estimate how these behavioral responses vary over the business cycle, proxied by the local unemployment rate, and we find suggestive evidence that the responses are lower in magnitude when the unemployment rate is relatively high, although our estimates are imprecisely estimated. We also find suggestive evidence that the ratio of the macro to micro participation responses decreases during times of high unemployment.

As an illustration, we use our empirical estimates to implement our sufficient statistics formula and calibrate the optimal income tax. We demonstrate three key results. First, relative to the optimal tax schedule in Saez (2002), we find that since the macro participation response is less than the micro response, this moves the optimal schedule more towards an NIT-like tax schedule with

[4] 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.

a relatively larger lump sum payment to the non-employed combined with higher employment tax rates. Second, we show that calibrating our tax formula with smaller 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. Third, we use our empirical estimates of behavioral responses over the business cycle to show that during recessions, the optimal income tax at the bottom shifts more towards an NIT-like structure.[5]

The primary advantage of our sufficient statistics approach is its generality with respect to the underlying mechanisms. In particular, competitive models with fixed and flexible wages (Diamond, 1980, Saez, 2002, 2004, Choné and Laroque, 2005, 2011, Rothstein, 2010, Lee and Saez, 2012) and models with matching frictions (Hungerbühler et al., 2006, Landais et al., 2016) are special cases of our sufficient statistics formula. To show the role of only allowing for flexible wages, we consider the competitive model with flexible wages by assuming that the conditional employment probability is either one (i.e., full employment) or does not respond to taxes (exogenous unemployment), and permit wages to respond to tax liabilities. Under the assumption that the production technology exhibits constant returns to scale (CRS) and workers are paid their marginal products, we show that the optimal tax formula exactly equals the tax formula in Saez (2002) where wages are fixed. Thus, only allowing for endogenous wages, but not endogenous unemployment, does not affect the optimal tax schedule. The other advantage of our tax formula is that it is exact and does not rely on any approximations. The disadvantage of our approach however is that analytical results about the precise shape of the optimal tax schedule are harder to obtain.

Our paper builds on and contributes to the literature on labor supply responses to taxation 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) and Leigh (2010) who consider labor demand and wage 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, which is necessary to implement our optimal tax formula, and we use a single methodology and the same sample.[6] Second, our results clarify the importance of distinguishing

[5] Interestingly, while governments have in general shifted away from NIT programs, in practice, transfers to the bottom are increased during recessions. For example, the U.S. significantly increased transfers to the non-employed through the Supplemental Nutrition Assistance Program (SNAP) during the Great Recession as part of the American Recovery and Reinvestment Act of 2009. This suggests that the shape of optimal income transfers at the bottom might depend on the strength of the labor market. Unfortunately, there is very little research on this question to help guide policymakers since current models by design do not allow for this possibility.

[6] 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. We estimate the micro elasticity using micro data and control for market fixed effects. For the macro elasticity, we pool the data to the market level and control separately for year and state fixed effects. One can show that this approach is essentially equivalent to one that estimates both the micro and macro equation in a single regression. This avoids the concern that differences in micro and macro

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.[7]

A number of recent papers have highlighted the distinction between micro and macro behavioral responses. The first paper to show that both are important for optimal policy is Landais et al. (2016), 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. We formally show that the optimal benefit level formula in Landais et al. (2016) is a special case of our model. In particular, we derive this formula under the following assumptions. First, we assume away cross effects so that wages or job-finding probabilities in one occupation do not respond to taxes in another occupation. Second, we assume that all of the labor supply responses occur along the search intensity margin, not the participation margin. Third, wage and tightness are assumed to depend on tax policy only through the difference in utility between employment and unemployment. Last, the social objective is assumed to be unweighted utilitarian. 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. (2016), some recent papers have tried to empirically estimate macro and micro effects of UI benefits (e.g. Lalive et al., 2015) and job search assistance programs (e.g. Crépon et al., 2013) on unemployment durations.[8]

The distinction between micro and macro responses also plays an important role in the recent literature estimating extensive and intensive labor supply responses (See Chetty et al., 2011a, and Chetty et al., 2012, for an overview). The terms micro and macro responses in these papers corre-

estimates are confounded by differences in methodologies and/or different samples.

[7]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.

[8] 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., 2016). Lalive et al. (2015) 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.

spond to conceptually the same responses that are identified using different sources of variation in taxes. 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., 2011b) 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 (2016) and Landais et al. (2016) 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 2016 and Landais et al. 2016) 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" reform in Norway are attenuated when the local unemployment rate is relatively high.

Finally, our work broadly relates to research which permit labor demand 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 and empirical results. Section IV considers the policy implications of our theoretical and empirical findings. The last section concludes.

II The theoretical model

This section derives optimal tax formulas. We first develop a framework that is consistent with a rich set of labor market responses to taxation (II.1). Following Chetty (2009), we use this benchmark model to identify the sufficient statistics that are necessary to compute the optimal income tax. Then, we specialize this model to connect our formula to previous formulas in the literature (II.2). Finally, we propose variants of our baseline model with more structure on the labor demand to extend for unemployment benefits different than welfare benefits (II.3) and for a tax on profits together with a continuous job search decision (II.4).

II.1 The benchmark model

We generalize 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 non-employment. All other occupations correspond to a specific labor market where the gross wage (equivalently pre-tax earnings) is w_i, the net wage (or consumption) is c_i and the tax liability is $T_i = w_i - c_i$.[9]. The timing is:

1. The government chooses the tax policy.

2. Each individual m chooses an occupation $i \in \{0, ..., I\}$ to participate in.[10]

3. For each labor market $i \in \{1, ..., I\}$, a fraction $p_i \in (0, 1]$ of participants are employed, receive gross wage w_i, pay tax T_i and consume the after-tax wage $c_i = w_i - T_i$. The remaining fraction $1 - p_i$ of participants are unemployed.

Unlike Saez (2002), we make a distinction among the *non-employed* individuals between the *unemployed* who search for a job in a specific labor market and fail to find one and the *non-participants* who choose not to search for a job.[11] 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 work, hereafter the *conditional employment probability*, and $h_i = k_i\, p_i$ denotes the number of employed workers.

[9]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

[10]Our model captures two types of labor supply responses along the intensive margin: moving between two consecutive occupations as in Saez (2002) and hours or in-work effort responses within a given occupation. The latter are captured through changes in earnings w_i.

[11]We simply assume job search intensity is either zero for non-participants or one for participants. In Section II.4 an extension with a continuous job search intensity and show continuous job search intensity *per se* does not change the optimal tax formula, but requires more structure on the labor market.

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 is $h_0 = k_0 + \sum_{i=1}^{I} k_i(1 - p_i)$.

We assume that all participants in a labor market i face the same probability p_i to be employed. This "uniform rationing" assumption is made for tractability, just as the assumption that all employed in a given labor market are paid the same wage w_i. All the non-employed, whether non-participants or unemployed, receive the same welfare benefit denoted b.[12] Therefore, the policy choice of the government is represented by the vector $\mathbf{t} = (T_1, ..., T_I, b)'$. The government faces the following budget constraint:

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

where $E \geq 0$ is an exogenous amount of public expenditures. One more employed worker in occupation i increases the government's revenues by the amount T_i of tax liability she pays, plus the amount of welfare benefit b she no longer receives, the sum of the two defining the *employment tax*.[13] 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.

Profits do not appear explicitly in our model. This is consistent with two possible scenarios. First, many natural models of the labor market, such as competitive models with constant returns to scale (Lee and Saez, 2012) or models with matching frictions on the labor market and free entry (Mortensen and Pissarides, 1999) have profits equal to zero in equilibrium. Second, our results are consistent with the presence of profits if we assume that profits are not taxed and if the welfare of capital owners who receive profits does not enter the social welfare function. These assumptions are clearly simplifying. We consider in subsection II.4 an extension of our model with partially taxed profits. Introducing a tax on profits requires us to place more structure on labor demand.

Micro vs. Macro Labor Supply Responses

In the paper, we make a crucial distinction between macro and micro responses to taxes. On the one hand, we define *micro* responses to a tax change in the hypothetical case where tax changes

[12] 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. In Section II.3, we will consider an extension of the model where unemployed in labor market i receive a benefit b_i that may differ from the welfare benefit z given to nonparticipants.

[13] 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.

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.

On the other hand, *macro* responses to tax policy $\mathbf{t}$ are defined to encapsulate the general equilibrium responses of wages and of conditional employment probabilities to taxes. To describe the latter, rather than specify the micro-foundations of the labor market, we use reduced-forms denoted $\mathscr{W}_i(\cdot)$, $\mathscr{C}_i(\cdot)$ and $\mathscr{P}_i(\cdot)$.[14] 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 is given 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. We only assume that these functions are differentiable, that $\mathscr{P}(\cdot)$ takes values in $(0,1]$ and that $0 < \mathscr{W}_1(\mathbf{t}) < ... < \mathscr{W}_I(\mathbf{t})$ and $b < \mathscr{C}_1(\mathbf{t}) < ... < \mathscr{C}_I(\mathbf{t})$ for all tax policies $\mathbf{t}$. The two latter assumptions ensure that occupations indexed with a higher i correspond to labor markets with higher before-tax and after-tax earnings.

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 for a job in labor market i, with the normalization $\chi_0(m) = 0$. Heterogeneity in $\chi_i(m)$ accounts for difference in the taste of work, but also for heterogeneity in skills. For instance, if individual m does not have the required skill to work in occupation i, it becomes extremely costly for her to apply to these jobs, in which case we consider that $\chi_i(m) = +\infty$.

Individual m enjoys a utility level equal to $u(c_i) - d_i - \chi_i(m)$ if she finds a job in labor market i, $u(b) - \chi_i(m)$ if she is unemployed in labor market i, and $u(b)$ if she chooses not to search for a job. Let $\mathscr{U}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathscr{P}_i(\mathbf{t}) \left(u\left(\mathscr{C}_i(\mathbf{t})\right) - d_i\right) + (1 - \mathscr{P}_i(\mathbf{t}))\, u(b)$ denote the *gross* expected utility of searching for a job in occupation i, absent any participation cost, as a function of the tax policy $\mathbf{t}$, and let:

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

denote its realization at a particular point of the tax system. Let $U_0 = u(b)$ be the utility expected out of the labor force. Individual m expects $U_i - \chi_i(m)$ 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 $j \in \{0, ..., I\} \setminus \{i\}$. Provided that the distribution of participation costs $(\chi_1, ..., \chi_I)$ is smooth enough, one can write the number k_i of participants in each labor market as a function denoted $\hat{\mathcal{K}}_i(\cdot)$ of gross expected utilities, so that:

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

[14] We here assume that an equilibrium exists and is unique. This equilibrium varies smoothly with the policy $\mathbf{t}$ in a way described by the $\mathscr{W}(\cdot)$, the $\mathscr{C}(\cdot)$ and the $\mathscr{P}(\cdot)$ functions.

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

$$\left.\frac{\partial \mathscr{U}_i}{\partial T_j}\right|^{\text{Micro}} = -p_i \, u'(c_i) \, \mathbb{1}_{i=j} \tag{3}$$

while the macro one is given by:

$$\frac{\partial \mathscr{U}_i}{\partial T_j} = \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{P}_i}{\partial T_j}\frac{u(c_i) - d_i - u(b)}{p_i \, u'(c_i)}\right] p_i \, u'(c_i) \tag{4}$$

The micro and macro responses of expected utilities to taxation differ for two reasons. First, for micro responses, unlike for macro ones, gross wages are held constant. Therefore, if $i = j$, the tax change is passed through one for one to the worker and $\frac{\partial \mathscr{C}_i}{\partial T_j} = -1$. If $i \neq j$, the tax reform has no effect on before-tax and after-tax wages, so $\frac{\partial \mathscr{C}_i}{\partial T_j} = 0$. Conversely, for macro responses, tax adjustments may affect wages in a variety of ways so in general, $\frac{\partial \mathscr{C}_j}{\partial T_j} \neq -\mathbb{1}_{i=j}$. Second, at the micro level, the conditional employment probabilities are unaffected by a change in taxes. Conversely, at the macro level, responses of conditional employment probabilities have to be taken into account. They may be due to changes in labor supply or due to changes in vacancy creation by employers, as we will discuss below. As a consequence, the matrix $\left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{\text{Micro}}$ of micro responses of expected utilities is diagonal while the matrix $\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}$ of macro responses may not be diagonal, in which case the tax liability in one occupation affects expected utility in another occupation.[15] We refer to the off-diagonal elements of the matrix of macro responses as "cross effects" throughout.

Applying the chain rule to (2) and using matrix notations, the macro and micro participation responses to taxation are given by $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} = \frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$ and $\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{\text{Micro}} = \left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{\text{Micro}} \cdot \frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$. This lead us to the following lemma:

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

$$\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathscr{U}}{\mathbf{dT}}\right|^{Micro}\right)^{-1} = \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{Micro}\right)^{-1} \tag{5}$$

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

The effect of a government policy on individual welfare is summarized by how that policy affects her gross expected utility. Since participation responses depend only on expected utilities, it follows that one may use the macro and micro participation responses to characterize how the effects of a policy on welfare differ at the macro level compared to the micro one.[16] We believe this "revealed preference" result is both important and novel. It suggests that labor force participation responses are a useful guide to understanding the welfare effects of government policies.[17]

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

$$h_i = \mathcal{H}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathcal{K}_i(\mathbf{t}) \mathscr{P}_i(\mathbf{t}) \tag{6}$$

Social objective

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

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

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

The optimal policy

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

$$\Lambda(\mathbf{t}) \stackrel{\text{def}}{\equiv} \sum_{i=1}^{I} (T_i + b) \, \mathcal{H}_i(\mathbf{t}) - b - E + \frac{1}{\lambda} \Omega \left(\mathscr{U}_1(\mathbf{t}), ..., \mathscr{U}_I(\mathbf{t}), u(b) \right) \tag{7}$$

[16] If matrix $\frac{\mathbf{d}\hat{\mathcal{K}}}{\mathbf{dU}}$ was not invertible, for instance if labor supply was exogenous in one occupation, then Equation (5) would be mathematically a nonsense.

[17] For example, Lavecchia (2016) considers a special case of our model to study optimal minimum wage policy. He obtains a revealed preference result that is very much in the same spirit as our result. In particular, he shows that the optimality of the minimum wage in the presence of an optimal non-linear income tax depends on the macro labor force participation response to the minimum wage.

Following Saez (2001, 2002) and Saez and Stantcheva (2016), we define the marginal social welfare weight of workers in occupation $i \in \{1, ..., I\}$ in terms of the micro response of the social objective to a tax change in occupation i, expressed in monetary terms:

$$g_i \overset{\text{def}}{\equiv} -\frac{1}{\lambda\, h_i} \left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \frac{\partial \Omega}{\partial U_i} \qquad \Leftrightarrow \qquad \frac{1}{\lambda}\frac{\partial \Omega}{\partial U_i} = -\left(\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \right)^{-1} g_i\, h_i \tag{8}$$

The social weight g_i represents the social value in monetary terms of transferring an additional dollar to an individual working in occupation i ignoring the responses of wages and of the conditional employment probabilities. It is non-negative according to (3). Absent these 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 (8), the first-order condition with respect to the tax liability T_j in labor market j is:

$$0 = \underbrace{h_j}_{\text{Mechanical effect}} + \underbrace{\sum_{i=1}^{I} (T_i + b)\, \frac{\partial \mathcal{H}_i}{\partial T_j}}_{\text{Behavioral effects}} - \underbrace{\sum_{i=1}^{I} \frac{\partial \mathscr{U}_i}{\partial T_j} \left(\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \right)^{-1} g_i\, h_i}_{\text{Social Welfare effects}} \tag{9}$$

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. This change incorporate general equilibrium (macro) responses. For each additional worker in occupation i, the government increases its resources by the employment tax $T_i + b$ that is equal to the additional tax received T_i plus the benefit b that is no longer paid.

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

on the social objective, the usual social welfare terms $g_i\ h_i$ have to be inflated by the ratio between the macro and the micro responses of expected utility to taxation.

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

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

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) - \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\left. \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \right|^{Micro} \right)^{-1} \cdot (\mathbf{g\ h}) \tag{10}$$

Proof: Equation (9) can be rewritten in matrix notations:

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

Using Equation (5) in Lemma 1 leads to (10). □

Due to the presence of wage and unemployment responses, our optimal tax formula differs from Saez (2002) in two important ways. First, the behavioral effects depend on the macro responses of employment and not on the micro labor supply responses. Second, as the welfare weights are computed from the micro effects of a tax reform on the social objective, while the macro responses are relevant to characterize the optimal tax, one needs a correcting procedure, which consists in multiplying in the optimal tax formula the (vector of) welfare weights times employment levels by the (matrix) ratio of macro to micro responses of expected utilities, as in Equation (11). From (4), estimating this ratio requires estimating the macro responses of wages to taxation, which is especially difficult to identify in practice. However, as participation decisions depends on taxation only through expected utilities, according to Lemma 1, the matrix ratio of macro over micro responses of expected utility to taxation coincides with the matrix ratio of macro over micro participation responses, which can be identified more easily, as will be illustrated in the empirical Section III.

Recently, several papers have derived optimal policy formulas that are expressed as the sum of two terms, a standard redistributive or insurance term which illustrates the equity-efficiency or insurance-incentives trade-off and a "correction" term which accounts for inefficiencies in the labor market.[18] To connect our optimal tax formula to the formulas derived in these papers, it is

[18] see Landais et al. (2016) for unemployment insurance, Fahri and Werning (2016) for optimal government spending and macroprudential policies and Michaillat and Saez (2017) for government spending.

useful to consider a "first-best environment"where the government directly controls wages and conditional employment probabilities. In such a case, the first-best Lagrangian writes:

$$\Lambda^1(\mathbf{t},\mathbf{w},\mathbf{p}) \stackrel{\text{def}}{\equiv} \sum_{i=1}^{I}(T_i+b)\,\mathcal{H}_i^1(\mathbf{t},\mathbf{w},\mathbf{p})$$
$$+\quad \frac{1}{\lambda}\Omega\left(p_1[u(w_1)-d_1]+(1-p_1)u(b),...,p_I[u(w_I)-d_I]+(1-p_I)u(b),u(b)\right)$$

where $\mathcal{H}_i^1(\mathbf{t},\mathbf{w},\mathbf{p})$ denotes the first-best level of employment in labor market i. In particular, employment responses to taxes in this environment correspond to micro employment responses which are solely driven by labor supply responses. Using (8), the optimal tax formula in the first-best case becomes:

$$\mathbf{0}=\frac{\partial\mathbf{\Lambda^1}}{\partial\mathbf{T}}\equiv(\mathbf{1}-\mathbf{g})\mathbf{h}+\left.\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\right|^{\text{Micro}}\cdot(\mathbf{T}+\mathbf{b})$$

which corresponds to Equation (11) in Saez (2002). As the two Lagrangians are related through $\Lambda(\mathbf{t})\ \equiv\ \Lambda^1\left(\mathbf{t},\mathscr{W}_1(\mathbf{t}),...,\mathscr{W}_I(\mathbf{t}),\mathscr{P}_1(\mathbf{t}),...\mathscr{P}_I(\mathbf{t})\right)$ the optimal tax formula (10) in Proposition 1 is equivalent to:

$$\mathbf{0}=(\mathbf{1}-\mathbf{g})\mathbf{h}+\left.\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\right|^{\text{Micro}}\cdot(\mathbf{T}+\mathbf{b})+\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}\cdot\frac{\partial\mathbf{\Lambda^1}}{\partial\mathbf{w}}+\frac{\mathbf{d}\mathscr{P}}{\mathbf{dT}}\cdot\frac{\partial\mathbf{\Lambda^1}}{\partial\mathbf{p}}$$

Therefore, our optimal tax formula coincides with Equation (11) in Saez (2002) if all wage and conditional employment probabilities are set at their first-best values for which $\frac{\partial\Lambda^1}{\partial w_j}=\frac{\partial\Lambda^1}{\partial p_j}=0$ for all $j\in\{1,...,I\}$.[19] Otherwise, optimal employment taxes are set away from their first-best optimal values to correct for the inefficient values of wages and conditional employment probabilities.

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):

$$0=-h_0+\sum_{i=1}^{I}(T_i+b)\frac{\partial\mathcal{H}_i}{\partial b}+g_0h_0+\sum_{i=1}^{I}g_ih_i\left[\frac{\partial\mathscr{C}_i}{\partial b}+\frac{\partial\mathscr{P}_i}{\partial b}\frac{u(c_i)-d_i-u(b)}{p_i\,u'(c_i)}\right] \tag{12}$$

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

$$g_0\stackrel{\text{def}}{\equiv}\frac{u'(b)}{\lambda\,h_0}\left[\frac{\partial\Omega}{\partial U_0}+\sum_{i=1}^{I}(1-p_i)\frac{\partial\Omega}{\partial U_i}\right] \tag{13}$$

In particular, if we furthermore assume there are 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 $\sum_{i=1}^{I}\frac{\partial\mathcal{H}_i}{\partial T_i}=\frac{\partial\mathcal{H}_i}{\partial b}$, the weighted sum of social welfare weights has to be equal to 1 (See Appendix A.1):

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

[19] These first-best values of wages and of conditional employment probabilities are not only based on efficiency considerations but depend also on insurance and redistribution issues as the terms $\frac{\partial\Omega}{\partial U_i}$ and $u'(c_i)$ appear in $\frac{\partial\Lambda^1}{\partial w_i}$ and in $\frac{\partial\Lambda^1}{\partial p_i}$.

II.2 Special cases of the optimal tax formula

In this section, we consider several special cases of our baseline model to connect our optimal tax formula to formulas in the public finance literature.

The no-cross effect case

To connect our optimal tax formula with the optimal tax formula in the pure extensive models of Diamond (1980) and Saez (2002), we first consider the "no-cross effect" case where: $\partial \mathscr{W}_i / \partial T_j = \partial \mathscr{C}_i / \partial T_j = \partial \mathscr{P}_i / \partial T_j = \partial \hat{\mathcal{K}}_i / \partial U_j = 0$ for $i \neq j$ and $i \neq 0$. This means that labor demand only responds to tax liabilities in the same market, but not in other markets. It also implies that labor supply responses are concentrated along the extensive margin; in other words, individuals can move from non-employment to work (or vice-versa) in a single occupation, but cannot move between occupations in response to a tax change. The no-cross effect assumption implies from (4) that $\partial \mathscr{U}_i / \partial T_j = 0$, thereby: $\partial \mathcal{K}_i / \partial T_j = \partial \mathcal{H}_i / \partial T_j = 0$ 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. Therefore, the matrices of behavioral responses in (10) are all diagonal under the no cross-effect assumption.

Let the micro participation elasticity be defined as $\pi_j^m \stackrel{\text{def}}{\equiv} -\frac{c_j - b}{k_j} \frac{\partial \mathcal{K}_j}{\partial T_j}\Big|^{\text{Micro}}$, the macro participation elasticity as $\pi_j \stackrel{\text{def}}{\equiv} \equiv -\frac{c_j - b}{k_j} \frac{\partial \mathcal{K}_j}{\partial T_j}$ and the macro employment elasticity as $\eta_j \stackrel{\text{def}}{\equiv} -\frac{c_j - b}{h_j} \frac{\partial \mathcal{H}_j}{\partial T_j}$. From (6), the macro *employment* response η_j verifies $\eta_j = \frac{c_j - b}{p_j} \frac{\partial \mathscr{P}_j}{\partial T_j} + \pi_j$. It includes conditional employment responses $\frac{c_j - b}{p_j} \frac{\partial \mathscr{P}_j}{\partial T_j}$ in addition to the macro *participation* responses π_j. Equation (10) then simplifies to:

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

$$\frac{T_j + b}{c_j - b} = \frac{1 - \dfrac{\pi_j}{\pi_j^m} g_j}{\eta_j} \tag{14}$$

The no-cross effect environment is the simplest one to understand how the introduction of unemployment and wage responses modifies the optimal tax formula. In the pure extensive case without unemployment, Diamond (1980), Saez (2002) and Choné and Laroque (2005, 2011) showed that the optimal tax formula takes the form of an inversed elasticity rule $\frac{T_j + b}{c_j - b} = \frac{1 - g_j}{\eta_j}$ which is in the spirit of Ramsey (1927). There are two key differences between the latter equation and Equation (14). First, the denominator in (14) corresponds to the macro employment elasticity. Second, equation (14) inflates the social marginal welfare weight by the ratio of the macro to micro participation elasticity. As explained in Lemma 1, this is to account for the macro effects of taxation on the social objective, since the welfare weights are defined in terms of the micro expected utility responses. To understand why, consider a decrease in tax liability T_j in the no-cross effects case. This triggers a positive direct impact on social welfare $-g_j h_j$, which is the only one at the micro level. Moreover, this decrease in tax liability typically induces a decrease in the gross wage when $\frac{\partial \mathscr{W}_j}{\partial T_j} > 0$, so the wage response attenuates the direct impact on social welfare. Finally, the decrease in tax liability also typically triggers a rise in job creation, 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 response dominates (is dominated by) the impact of the wage response. Proposition 2 implies:

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

According to (14), a negative employment tax (EITC) becomes optimal whenever the social welfare weight is higher than the ratio of the micro over the macro participation elasticity, instead of one without unemployment and wage responses. Estimating the ratio of macro over micro participation elasticity is therefore sufficient to conclude whether unemployment and wages responses makes the EITC more or less likely optimal in the case where there are no cross effects.

The matching model with linear technology and proportional bargaining

We now consider the case of a Diamond-Mortensen-Pissarides (DMP) search-and-matching model with a linear production function and proportional bargaining. Following (Diamond, 1982, Pissarides, 1985, Mortensen and Pissarides, 1999, Pissarides, 2000) we assume a constant returns to scale matching function specific to each labor market which provides 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 it more rewarding for firms to hire a worker) and the more numerous job-seekers there are (which decreases the search congestions from the firm's 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.[20] 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, then 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

[20] We derive in Appendix A.2 this standard result, as well as the proof of Proposition 3 below.

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:[21]

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

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 $\beta_i = \mu_i$ 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 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, the gross wage 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 2 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.[22] Therefore, in the DMP model the macro-micro participation gap

[21] A similar expression for wage bargaining appears in Jacquet et al. (2014) and in Landais et al. (2016).

[22] As $\frac{\pi_j}{\pi_j^m} = \frac{\beta_j}{\mu_j}$ from (30), Equation (14) 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).

It is worth noting that under the Hosios (1990) condition $\beta_i = \mu_i$, while the macro and the micro *participation* elasticities are equal, this does not imply that the macro *employment* elasticities is equal to the micro *employment* elasticity. At the micro level, for fixed wages and tightness, a 1% increase in tax reduces employment only through the reduction in participation. The micro employment elasticity is therefore equal to the micro participation elasticity. Under the Hosios can be higher or lower than one, attenuating or reinforcing the arguments in favor of a negative participation tax at the bottom.[23]

The case with exogenous unemployment rates, constant returns to scale technology and no search costs

Saez (2004) argues that in a competitive model with a constant returns to scale technology and no matching frictions, the optimal tax formula is identical to one with fixed wage rates. We now relate our optimal tax formula to this result. For this purpose, we consider a model with exogenous unemployment rates in each labor market, a constant returns to scale technology and flexible wages that adjust to marginal products of labor. Under perfect substitution across the different types of labor, labor demand is perfectly elastic, wages are unresponsive to taxation, so (10) reduces immediately to Saez (2002, Equation (11)). If conversely the different types of labor are imperfect substitutes, wages may respond to tax reforms, so the micro and macro participation responses may be different due to the wage adjustments in each labor market. However Saez (2004) showed that in such a model without unemployment, the optimal tax formula can be expressed using only the micro employment response and takes the same form as Saez (2002) because the adjustment of wages compensate each other along the zero-profit price frontier. We show in Appendix A.3 that this "Production Efficiency"argument remains valid if unemployment rates are positive but exogenous.

Proposition 4. *In a model without search costs, with a constant return to scale technology and exogenous unemployment rates, if* $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ *is invertible, the optimal tax formula* (10) *is equivalent to:*

$$\forall j \in \{1, ..., I\} \qquad 0 = (1 - g_j)h_j + \sum_{i=1}^{I} (T_i + b) \left. \frac{\partial \mathcal{H}_i}{\partial T_j} \right|^{Micro} \tag{16}$$

This result shows that only endogenous unemployment changes the tax formula in Saez (2002). In particular, with wage responses and exogenous unemployment, the formula in Saez (2002) remains valid if wages adjust to the marginal product of labor and the technology exhibits constant returns to scale.[24]

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

(1990) condition, the latter is equal to the macro participation elasticity. However, as a 1% increase in tax also decreases tightness because of the wage response to taxes, the conditional employment probability is also reduced, so the macro employment response is larger than the macro participation response.

[23]By extending this model with intensive labor supply decision, the present model can include the central mechanism of Golosov et al. (2013) where firms have different productivity and individuals direct their search.

[24]In the no-cross effect case, matrix $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ would no be invertible if for one occupation, the tax incidence falls entirely on firms. The technical assumption that matrix $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ is invertible is therefore not very restrictive in practice.

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 in models that allowed for wages being set endogenously above the market clearing level (McDonald and Solow, 1981, Shapiro and Stiglitz, 1984, Akerlof and Yellen, 1990).[25]

To develop some intuition about the macro-micro participation gap in job-rationing models, we now consider a model with labor supply responses concentrated along the extensive margin, 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.[26] We assume that individuals who participate face a heterogeneous participation 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 em-

[25] 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.

[26] Note that with a fixed wage, it is no longer equivalent whether the firm or the worker pays the tax. If the firm pays 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 it optimal to set the cost of labor above the market-clearing level, thereby generating unemployment in a job-rationing model.

ployment 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. (2016) 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.

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 the tax schedule becomes more progressive, the wedge between net and gross wages increases more rapidly with the wage. 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, which may increase employment.[27] 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.

To understand how the wage moderating effects of tax progressivity affects the optimal tax schedule, let us consider a matching model with two occupations $I = 2$ and a linear production function, so that the conditional employment probability in one labor market is a decreasing function $p_i = \mathscr{L}_i(w_i)$ of the gross wage on that labor market. Assume that labor supply responses are concentrated along the extensive margin. Assume that the wage functions $\mathscr{W}_i$ not only verify $\frac{\partial \mathscr{W}_i}{\partial T_i} > 0$ for $i = 1, 2$, 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 that $\frac{\partial \mathscr{W}_2}{\partial T_1} \geq 0 \geq \frac{\partial \mathscr{W}_1}{\partial T_2}$, with at least one strict inequality. Then we have $\frac{\partial \mathscr{P}_i}{\partial T_i} > 0$ and $\frac{\partial \mathcal{H}_i}{\partial T_i} < 0$ for $i = 1, 2$ and $\frac{\partial \mathscr{P}_2}{\partial T_1} \leq 0 \leq \frac{\partial \mathscr{P}_1}{\partial T_2}$ with at least one strict inequality, thereby $\frac{\partial \mathcal{H}_2}{\partial T_1} < 0 < \frac{\partial \mathcal{H}_1}{\partial T_2}$, with at

[27] 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) or in the competitive directed search model (or wage posting) of Moen (1997) by Lehmann et al. (2011). 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. (2016) provide some empirical support for the unemployment reducing effect of tax progressivity.

least one strict inequality. Suppose that only one of the latter inequalities is strict and that only the welfare of the non employed is valued, so that $g_0 > g_1 = g_2 = 0$. For notational compactness, we assume $\frac{\partial \mathcal{H}_i}{\partial T_j} < 0 = \frac{\partial \mathcal{H}_j}{\partial T_i}$ with $j \in \{1,2\} \setminus \{i\}$. Inverting (10) leads to:

$$T_i + b = -\frac{h_i}{\frac{\partial \mathcal{H}_i}{\partial T_i}} > 0 \qquad \text{and} \qquad T_j + b = -\underbrace{\frac{h_j}{\frac{\partial \mathcal{H}_j}{\partial T_j}}}_{>0} + \frac{\partial \mathcal{H}_i}{\partial T_j} \underbrace{\frac{h_i}{\frac{\partial \mathcal{H}_i}{\partial T_i} \frac{\partial \mathcal{H}_j}{\partial T_j}}}_{>0}$$

Therefore, when $\frac{\partial \mathscr{W}_1}{\partial T_2} = 0 < \frac{\partial \mathscr{W}_2}{\partial T_1}$, so that $\frac{\partial \mathcal{H}_1}{\partial T_2} = 0 > \frac{\partial \mathcal{H}_2}{\partial T_1}$, the wage moderating effect of tax progressivity captured by the negative term $\frac{\partial \mathcal{H}_2}{\partial T_1}$ has no effect on the optimal tax for the high-skilled and tends to reduce the optimal tax for the low-skilled. Conversely when $\frac{\partial \mathscr{W}_1}{\partial T_2} < 0 = \frac{\partial \mathscr{W}_2}{\partial T_1}$, so that $\frac{\partial \mathcal{H}_1}{\partial T_2} > 0 = \frac{\partial \mathcal{H}_2}{\partial T_1}$, the wage moderating effect of tax progressivity captured by the positive term $\frac{\partial \mathcal{H}_2}{\partial T_1}$ has no effect on the optimal tax for low-skilled and tends to increase the optimal tax for the high-skilled. In these two simplistic cases, we retrieve the general result shown by Hungerbühler et al. (2006), Lehmann et al. (2011) in more specialized search matching models, that 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.

II.3 Different unemployment and welfare benefits

We now extend our baseline model to allow for different benefits for the unemployed and the non participants. To differentiate the benefits provided to the non-employed, the government needs to perfectly observe the participation decisions of individuals. In such a case, the government provides non-participants with welfare benefits z and provides unemployed in labor market i with unemployment benefits b_i. The policy vector therefore becomes $\mathbf{t} = (T_1, ..., T_i, b_1, ..., b_I, z)$ and the budget constraint (1) becomes:[28]

$$\sum_{i=1}^{n} T_i\, h_i = z\, k_0 + \sum_{i=1}^{n} (k_i - h_i) b_i + E \qquad \Leftrightarrow \qquad \sum_{i=1}^{n} (T_i + b_i)\, h_i + \sum_{i=1}^{n} (z - b_i) k_i = z + E \tag{17}$$

For a fixed participation level k_i in labor market i, an additional employed worker pays tax T_i and no longer receives unemployment benefits b_i. Therefore, for each additional employed worker in labor market i holding the number of participants constant, the government's revenue increases by the *employment tax* $T_i + b_i$. Symmetrically, for a fixed employment level h_i in labor market i, an additional participant receives the unemployment benefits b_i instead of the welfare benefits z. Therefore, the government's revenue changes by $z - b_i$ for each additional participant in labor market i holding the level of employment constant. The Lagrangian (7) of the government's

[28]We here use that $k_0 = 1 - \sum_{i=1}^{I} k_i$.

problem thus becomes:

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

Using (8), the first-order condition with respect to tax liability T_j in labor market j becomes:

$$0 = h_j + \sum_{i=1}^{I} (T_i + b_i)\frac{\partial \mathcal{H}_i}{\partial T_i} + \sum_{i=1}^{I} (z - b_i)\frac{\partial \mathcal{K}_i}{\partial T_i} - \sum_{i=1}^{I} \frac{\partial \mathscr{U}_i}{\partial T_j}\left(\left.\frac{\partial \mathscr{U}_i}{\partial T_j}\right|^{\text{Micro}}\right)^{-1} g_i h_i$$

Compared to (9), different benefits for the unemployed and the non-participants generates a new fiscal externality: for a fixed level h_i of employment in labor market i, each additional participant increases the number of unemployment benefits b_i recipients and reduces the number of welfare benefits recipients z. Moreover, the employment tax now depends on the unemployment benefits b_i. As Lemma 1 continues to hold, the optimal tax formula (10) becomes:

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} \cdot (\mathbf{T} + \mathbf{b}) + \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot (\mathbf{z} - \mathbf{b}) - \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{\text{Micro}}\right)^{\mathbf{-1}} \cdot (\mathbf{g}\ \mathbf{h}) \tag{18}$$

Therefore, the same sufficient statistics, namely the macro employment responses $\frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}$, the macro participation responses $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}$ and the micro participation responses $\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{\text{Micro}}$ have to be estimated to implement the optimal tax formula. Clearly, when unemployment and welfare benefits are equal, we retrieve formula (10).[29]

II.4 The introduction of profits and endogenous search intensity

Up to now, profits did not appear in our model. We assumed that if firms make profits, these profits are untaxed and these profits are received by some ”capital owners” whose welfare is not included in the social welfare function. Alternatively, the public finance literature has considered a polar assumption where profits are fully taxed, or, equivalently, where all production is controlled by the government (Diamond and Mirrlees, 1971). It is therefore important to consider an extension of our model where profits are taxed at an exogenous rate denoted $\tau \in [0, 1]$. Our baseline model corresponds to the case where $\tau = 0$ while the case of fully taxed profits corresponds to $\tau = 1$.

To introduce a tax on profits, we need to specify how profits appear, so we have to specify the production technology and the matching technology. We consider a model where a representative firm produces a numeraire good using the different types of labor under the technology $F(h_1, ..., h_I)$ which is increasing and weakly concave in each of its I arguments so $F_i > 0 \geq F_{ii}$. We assume that creating a vacancy costs $\kappa_i > 0$ to the firm.

[29] In Appendix A.4, we derive the formulas for the optimal unemployment and welfare benefit levels.

To have a general model that includes Landais et al. (2016) as a special case, we introduce a continuous search intensity denoted e. Firms open ν_i vacancies on labor market i, while the total amount of search units provided by the k_i participants is denoted S_i. In particular, in equilibrium, all participants choose the same amount of search units, in which case $S_i = e_i k_i$. The employment level h_i in labor market i is given by the matching function $\mathcal{M}_i(\nu_i, S_i)$ which is increasing in each of its two arguments and exhibits constant returns to scale. Let $\theta_i \stackrel{\text{def}}{\equiv} \nu_i / S_i$ be the "vacancy search units ratio". Each vacancy is filled with a probability $q_i = \mathcal{M}_i(\nu_i, S_i)/\nu_i = \mathcal{M}_i(1, 1/\theta_i) = Q_i(\theta_i)$, where $Q_i(\cdot)$ is decreasing in the vacancy search units ratio. Symmetrically, the probability of finding a job per unit of search in labor market i is $a_i = \mathcal{M}_i(\nu_i, S_i)/S_i = \mathcal{M}_i(\theta_i, 1) = P_i(\theta_i)$, where $P_i(\cdot)$ is increasing in the vacancy search units ratio. We define tightness in labor market i by the probability a_i of finding a job per unit of search.[30] The conditional probability to find a job is $p_i = e_i a_i$.

To hire one more worker, the firm has to post $1/Q_i(\theta_i)$ vacancies, which costs $\kappa_i / Q_i(\theta_i)$. Let $J_i(a) \stackrel{\text{def}}{\equiv} \kappa_i / Q_i\left(P_i^{-1}(a)\right)$ denote the hiring cost for the firm as an increasing function of tightness a and let $\mathscr{A}_i(\mathbf{t})$ be the reduced form describing how the tax policy affects tightness a_i in labor market i at the general equilibrium. The representative firm chooses labor demand $h_1, ..., h_I$ to maximize profits, taking wages $\mathbf{w} = (w_1, ..., w_I)$ and tightnesses $\mathbf{a} = (a_1, ..., a_I)$ as given:

$$\Pi(\mathbf{t}) \stackrel{\text{def}}{\equiv} \max_{h_1,...,h_I} \qquad F(h_1, ..., h_I) - \sum_{i=1}^{I} \left(\mathscr{W}_i(\mathbf{t}) + J_i\left(\mathscr{A}_i(\mathbf{t})\right)\right) h_i \tag{19}$$

The labor demand first-order conditions are:

$$F_i(h_1, ..., h_I) = w_i + J_i(a_i) \qquad \forall i \in \{1, ..., I\}$$

For a participant in labor market i, searching for a job with intensity e induces a search cost equal to $D_i(e)$ where function $D_i(\cdot)$ is assumed increasing and convex. Each participant chooses search intensity taking the wage w_i and the tightness a_i as given. Individual m expects $U_i - \chi_i(m)$ by searching a job in labor market i where:

$$U_i \stackrel{\text{def}}{\equiv} \max_{e} \qquad e\, a_i\, \left(u(w_i - T_i) - d_i\right) + (1 - e\, a_i)\, u(b) - D_i(e) \tag{20}$$

Let:

$$\mathscr{U}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \max_{e} \qquad e\mathscr{A}_i(\mathbf{t})\, \left(u\left(\mathscr{C}_i(\mathbf{t})\right) - d_i\right) + \left(1 - e\mathscr{A}_i(\mathbf{t})\right) u(b) - D_i(e)$$

According to Equation (2), the number of participants in labor market i depends on taxation only through the responses of gross expected utility U_i to taxation. Therefore, our revealed preference

[30]It is usual in the matching literature to define instead tightness as the vacancy search units ration. However, as there is a one-to-one increasing relation between the two, there is no loss of generality in choosing either of the two definitions.

argument still applies and Lemma 1 continues to hold. With the additional tax revenues from profits, and assuming the same benefits for all the non-employed to save on notations, the Lagrangian (7) becomes:

$$\Lambda(\mathbf{t}) \stackrel{\text{def}}{\equiv} (T_i + b)\mathcal{H}_i(\mathbf{t}) + \tau\,\Pi(\mathbf{t}) + \frac{1}{\lambda}\Omega\left(\mathscr{U}_1(\mathbf{t}), ..., \mathscr{U}_I(\mathbf{t}), u(b)\right)$$

Using Hotelling's lemma[31] and Equation (8), the condition for the optimal tax liability T_j is:

$$0 = h_j + \sum_{i=1}^{I}(T_i + b)\frac{\partial \mathcal{H}_i}{\partial T_j} - \tau\sum_{i=1}^{I}\left(\frac{\partial \mathscr{W}_i}{\partial T_j} + J_i'(a_i)\frac{\partial \mathscr{A}_i}{\partial T_j}\right)h_i - \sum_{i=1}^{I}\frac{\partial \mathscr{U}_i}{\partial T_j}\left(\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}}\right)^{-1} g_i h_i \tag{21}$$

Using matrix notations and Equation (5) in Lemma 1, this condition becomes:

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\cdot(\mathbf{T}+\mathbf{b}) - \tau\left[\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}\cdot\mathbf{h} + \frac{\mathbf{d}\mathscr{A}}{\mathbf{dT}}\cdot(\mathbf{J_i'}(\mathbf{a_i})\ \mathbf{h_i})\right] - \frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\cdot\left(\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}}\right|^{\text{Micro}}\right)^{-\mathbf{1}}\cdot(\mathbf{g}\ \mathbf{h}) \tag{22}$$

A first striking feature of (22) is that search intensity responses do not appear explicitly. This is because the way taxation affects search intensity is very similar to the way taxation affect participation decisions, as shown in Appendix A.5. Consequently, behavioral effects in terms of search intensity are encapsulated in the macro employment responses. Moreover, provided that $\frac{\mathbf{d}\mathscr{E}}{\mathbf{dT}}^{\text{Micro}}$ is invertible, which is not the case when search intensity is exogenous, the optimal tax formula (22) can be equivalently expressed as follows:

$$\mathbf{0} = \mathbf{h} + \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}}\cdot(\mathbf{T}+\mathbf{b}) - \tau\left[\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}\cdot\mathbf{h} + \frac{\mathbf{d}\mathscr{A}}{\mathbf{dT}}\cdot(\mathbf{J_i'}(\mathbf{a_i})\ \mathbf{h_i})\right] - \frac{\mathbf{d}\mathscr{E}}{\mathbf{dT}}\cdot\left(\left.\frac{\mathbf{d}\mathscr{E}}{\mathbf{dT}}\right|^{\text{Micro}}\right)^{-\mathbf{1}}\cdot(\mathbf{g}\ \mathbf{h}) \tag{23}$$

This is related to results in Chetty (2008) who shows that one can use search effort behavioral responses to unemployment benefits to value Unemployment Insurance. However, since search intensity is typically unobserved, $\frac{d\mathscr{E}}{dT}$ and $\frac{d\mathscr{E}}{dT}^{\text{Micro}}$ cannot be estimated unless one imposes a normalization.[32]

The second noteworthy feature of Equation (22) is that compared to (9) or (10), several new terms are present when profits are taxed. This is because a change in tax on labor of type j affects wages w_i and recruitment costs $J_i(a_i)$, which triggers a change in the profit tax base. The optimal tax formula (22) shows that one needs to additionally account for the macro response of wages to taxation $\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}$, and market tightness, $\frac{\mathbf{d}\mathscr{A}}{\mathbf{dT}}$. Compared to labor supply responses, it is more difficult to identify these responses. For example, one needs to account for selection effects when trying to estimate the wage response to taxes.

There are two cases where these additional fiscal spillover responses do not appear in the optimal tax formula (22). The first case is when the tax on profits iz zero ($\tau = 0$) in which case

[31] That is, applying the envelope theorem to (19) implies that: $\frac{\partial \Pi}{\partial T_j} = -\sum_{i=1}^{I}\left(\frac{\partial \mathscr{W}_i}{\partial T_j} + J_i(a_i)\frac{\partial \mathscr{A}_i}{\partial T_j}\right)h_i$.

[32] Chetty (2008) assumes that search effort is equal to the hazard rate out of unemployment, which can be estimated using labor market flows.

we retrieve formula (10). The second case is when profits are fully taxed ($\tau = 1$) and when the behavioral responses to tax reforms are parameterized with respect to changes in utility levels $\mathbf{\Delta} = (\Delta_1 = u(c_1) - u(b), ..., \Delta_I = u(c_I) - u(b)$ instead of being parameterized with respect to changes in tax liabilities $(T_1, ..., T_I)$. In such a case, the macro responses of wages to taxation disappear from the optimal tax formula because the resource constraint depends only on employment and on after-tax incomes, pre-tax incomes being absent. This argument was made in Landais et al. (2016). The wage responses that show up when behavioral responses are parameterized with respect to changes in tax liabilities are now encapsulated in the behavioral responses parameterized with respect to changes in after-tax income. In Appendix A.6, we show that if *i*) there are no cross-effects and $I = 1$, *ii*) all the labor supply responses occur along the search intensity margin, not the participation margin, *iii*) wage and tightness depends on tax policy only through the difference in utility between employment and unemployment and *iv*) the social objective is unweighted utilitarian, we may recover the optimal tax formula of Landais et al. (2016) from our formula (22) as follows:

$$R \quad = \quad \frac{h}{\varepsilon^{\text{Micro}}} \frac{\Delta}{w} \left[\frac{1}{u'(c)} - \frac{1}{u'(b)} \right] + \left[1 - \frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}} \right] \frac{1}{1 + \varepsilon^{\text{f}}} \left[\frac{\Delta}{w \, \lambda} + (1 + \varepsilon^{\text{f}}) R - \frac{a \, J'(a)}{w} \right]$$

where $\Delta(T, b) \stackrel{\text{def}}{\equiv} u\left(\mathscr{C}(T, b)\right) - u(b) = u\left(\mathscr{W}(T, b) - T\right) - u(b)$, $R \stackrel{\text{def}}{\equiv} \frac{T+b}{w}$, $\varepsilon^{\text{Micro}} \stackrel{\text{def}}{\equiv} \frac{\Delta}{1-h} \left.\frac{\partial \hat{\mathcal{H}}}{\partial \Delta}\right|^{\text{Micro}}$, $\varepsilon^{\text{Micro}} \stackrel{\text{def}}{\equiv} \frac{\Delta}{1-h} \frac{\partial \hat{\mathcal{H}}}{\partial \Delta}$, and $\varepsilon^{\text{f}} \stackrel{\text{def}}{\equiv} \frac{a}{e} \frac{\partial e}{\partial a}$. This formula differs from Equation (23) of Landais et al. (2016) only by the term $\dfrac{a \, J'(a)}{w}$ instead of $\frac{\eta}{1-\eta} \tau(\theta)$ in Landais et al. (2016). Both terms corresponds to recruiting costs for the firm. However, in our model (as in most of matching models), recruiting costs are cost per vacancy posted. In Landais et al. (2016), these costs correspond to the hiring cost of recruiting workers. So both are conceptually identical although slightly different in the details. As policymaking is typically more about how much tax different types of individuals should pay rather than how much after-tax income they should receive, it seems much more natural to us to parameterize behavioral responses with respect to tax liabilities.

III Estimating Sufficient Statistics

To illustrate the practical relevance of our optimal tax formula, 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 single men and 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 individuals age 18 to 55 who are not in the military or enrolled full time in school or college. Since there is insufficient tax variation for higher income individuals over our sample period, we further restrict the sample to individuals with education less than a bachelors degree. Our theory distinguishes between individuals who choose to participate in the labor force (and are employed or unemployed) and those individuals who are actually employed. We measure these labor market states using the standard International Labor Office (ILO) criteria. A person is classified as being in the labor force if she is either employed or unemployed (i.e., actively looking for a job during the reference week and was available for work) and employed if she has been working during the reference week (or been temporarily absent from a job).[33]

Panel A of Table 1 shows descriptive statistics for the demographic characteristics of single individuals in the March CPS for the full sample (Column 1) and broken down by educational attainment groups (Columns 2-4), pooling all years from 1984 to 2011.[34] The age range is pretty similar across the three education groups - less than high school, high school, and some college - but there are large differences in the distribution of number of children, with lower educated single individuals being more likely to be parents. This is likely due to our sample restriction to singles since higher educated parents are more likely to be married. Additionally, low educated individuals are more likely to be black or Hispanic than high educated ones. Panel B displays labor market variables by educational attainment. Lower educated individuals are much less likely to be in the labor force than higher educated ones and also experience higher unemployment rates.

Tax and Transfer Calculator

In order to estimate the employment and participation effects of taxes and transfers it is necessary to compute the budget sets that individuals face. For this purpose, we developed a calculator that computes taxes and transfers at (nominal) income levels for single men and women, depending on the number of children, state and year.[35] We assume that an individual is filing as the head of the household and is claiming his/her children as dependents. To compute taxes (covering

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

[34] 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.

[35] We describe in details below how we impute income that serves as an input to the calculator.

federal and state income taxes, including tax credits, as well as FICA liability), we rely on the NBER TAXSIM software. We assign taxes based on state of residence, as reported in the CPS, as well as number of children, year, and income.[36] To compute transfers, in particular Aid to Families with Dependent Children (AFDC), Temporary Assistance for Needy Families (TANF) and Supplemental Nutrition Assistance Program (SNAP), we construct a benefit calculator based on rules published in the Welfare Rules Database, managed by the Urban Institute. This allows us to compute the benefits an individual is eligible for, as a function of number of children, state of residence, year and income. The shift from AFDC to TANF introduced a number of additional work and eligibility requirements for welfare recipients. For example, federal rules require a minimum number of TANF recipients to be employed and the lifetime duration of receiving TANF benefits is limited to a total of 5 years.[37] Rather than incorporate all of these policies explicitly into our empirical framework, we multiply benefits by gender specific recipiency rates constructed from the Survey of Income and Program Participation (SIPP). The new eligibility requirements are reflected in lower observed recipiency rates in our sample post-welfare reform.

We use our tax and transfer calculator to compute the incentive to work. Since we focus solely on the extensive margin in our analysis, we capture work incentives using just two measures, the transfer an individual receives when she has zero income and the tax and transfer level at the earnings level an individual obtains when working. A key difficulty is that earnings, and hence tax liabilities, are unobserved for non-employed individuals. Moreover, earnings for employed workers may be endogenous to the tax system. We proceed using two approaches. First, we impute an individual's tax liability following the approach taken in Eissa and Hoynes (2004) and Gelber and Mitchell (2012). We 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 ($X_{m,e,s,t}$): [38]

$$log(w_{m,e,s,t}) = \delta_{e,s,t} + X_{m,e,s,t}\pi_{e,t} + \epsilon_{m,e,s,t} \tag{24}$$

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

[36]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.

[37]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.

[38]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.

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.[39] We then use predicted earnings to impute an individual's tax liability using TAXSIM and the benefit calculator described above.

In the Online Appendix, we present OLS regressions of participation and employment using this imputed tax liability. One problem with this approach is that the demographic distribution itself, and therefore the imputed tax liabilities, might be endogenous to tax policy. For instance, more generous transfers to singles with kids, but not to individuals without children, may boost fertility and impact earnings. To address this concern, we also rely on a simulated instrument approach based on Currie and Gruber (1996).[40] This approach isolates policy variation in tax liabilities since it uses a fixed income and demographic distribution 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 employed individual in the sample. Second, using earnings for the full sample of employed 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 across all states and years. 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 of children, with variation driven solely by exogenous changes in the tax code, and not by endogenous changes in the earnings and/or demographics 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.

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

[40] 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.

Panel C of Table 1 shows the mean imputed real earnings for each education group averaged over the years and the corresponding tax and transfer levels depending on the number of children in the household. All numbers are reported in real 2010 U.S. dollars (USD). For high school dropouts, taxes (transfers) are strongly decreasing (increasing) in the number of children. 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 recipiency rates which are much less than 100 percent during our sample period. The last four rows report recipiency rates, as estimated in the SIPP. Each individual in the CPS is assigned a recipiency rate that we calculate from the SIPP based on gender, education, income and year. The table reports the average of the assigned recipiency rates separately for AFDC/TANF and food stamps, and also pre- and post-1996. We see that for high school dropouts, recipiency rates are roughly 50 percent for AFDC/TANF but fall to 20 percent post-1996. For food stamps, recipiency rates are much more comparable pre- and post-1996 and equal to roughly 40 percent.[41] These recipiency rates decrease with education which reflects diminishing eligibility as earnings increase.

III.2 Empirical Method

Specification of Labor Markets

In the theoretical model, individuals sort themselves into $I+1$ distinct occupations. 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), state and time (year-month). 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).

Estimating Micro and Macro Participation Responses and Macro Employment Responses

Equation (10) shows that the optimal tax schedule can be expressed in terms of macro employment responses and the ratio of macro to micro participation responses in matrix terms. 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 esti-

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

mation, we focus on the no-cross effects case where the above mentioned matrices are diagonal. We also assume away income effects by estimating the responses to employment tax liabilities $T_i + b$, instead of estimating separately the responses to tax liability T_i and to benefit b.

In our model $\mathcal{H}_i$ and $\mathcal{K}_i$ correspond to the number of individuals in income group i, but for an empirical specification that uses variation across individuals and labor markets, it makes little sense to assume $\frac{\partial \mathcal{K}_i}{\partial T_i}$ or $\frac{\partial \mathcal{H}_i}{\partial T_i}$ are constant across labor markets. Instead we will estimate the effect of taxes T_i on employment and participation *rates*. We denote the employment rate in income group i, which in our empirical setting will correspond to an education group i, as $\hat{\mathcal{H}}_i$ and the participation rate as $\hat{\mathcal{K}}_i$. These are the fraction of individuals with education level i who are employed or participating in the labor force, respectively. Estimating the marginal effects of taxes on employment and participation rates furthermore has the important advantage that the estimates are easier to interpret and to compare to the prior literature. For example, these estimates are straightforward to convert to employment and participation elasticities.

To obtain an econometric specification for the responses to taxation that is motivated by the theoretical model (without cross effects), we make two assumptions. First, we assume that the conditional employment probability and wage in a market can be written as functions of the average tax liability in that market only.[42] Second, we assume that tax liabilities vary across individuals within a labor market according to the number n of children in the household.[43] The function describing participation decisions for individual m in labor market (e, s, t) can thus be written as:

$$\hat{\mathcal{K}}_{m,e,s,t,n}(\mathbf{t}) = \tilde{\mathcal{K}}_{m,e,s,t,n}(p_{e,s,t}(T_{e,s,t}), w_{e,s,t}(T_{e,s,t}), T_{e,s,t,n}) \tag{25}$$

To estimate the micro participation response, we take a linear approximation to Equation (25), add labor market fixed effects (one FE for each state-by-year-by-month-by-education cell) and flexible controls (education by number of children FE, and demographic control variables like age, age-squared, race, ethnicity all interacted with education groups), to get the following econometric specification:

$$\hat{k}_{m,e,s,t,n} = T_{e,s,t,n}\beta + \delta_{e,s,t} + \delta_{e,n} + X_{m,e,s,t,n}\lambda + \nu_{m,e,s,t,n} \tag{26}$$

This equation implies that $\beta = \frac{\partial \hat{\mathcal{K}}_{m,e,s,t,n}}{\partial T_{e,s,t,n}}\Big|^{\text{Micro}}$ captures the micro participation effect. Implicit in this specification is a pooling assumption, whereby the partial derivative of taxes on participation does not vary across labor markets. We adopt this assumption for simplicity and because it is difficult to generate exogenous variation in tax liabilities that differentially affects income groups.

Next, to estimate macro participation responses, we aggregate the data to state-year-education averages, add education-by-year and education-by state fixed effects, region specific linear time

[42]The restrictions on the econometric specification correspond to the no-cross effect theoretical assumption that is assumed in Proposition 1 and Corollary 1.

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

trends, and demographic controls (cell averages of the micro controls) interacted with education to get:

$$\hat{k}_{e,s,t} = T_{e,s,t}\gamma + \delta_{e,s} + \delta_{e,t} + X_{e,s,t}\lambda + \nu_{e,s,t} \tag{27}$$

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. Therefore the macro effect can be obtained as: $\gamma = \frac{d\hat{\mathcal{K}}_{e,s,t}}{dT_{e,s,t}}$.[44]

The market-level employment rate in market (e, s, t) is given by $\hat{\mathcal{H}}_{e,s,t}(T_{e,s,t}) = p_{e,s,t}(T_{e,s,t}) \times \hat{\mathcal{K}}_{e,s,t}(T_{e,s,t})$. Thus, the macro employment response is given by $\frac{\partial \hat{\mathcal{H}}_{e,s,t}}{\partial T_{e,s,t}} = p_{e,s,t} \times \frac{\partial \hat{\mathcal{K}}_{e,s,t}}{\partial T_{e,s,t}} + \hat{\mathcal{K}}_{e,s,t} \times \frac{\partial p_{e,s,t}}{\partial T_{e,s,t}}$. We will rely on a linear approximation for the market-level employment rate similar to Equation (27) 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 parameter β, we require that the micro tax liability $T_{e,s,t,n}$ is exogenous, conditional on labor market and education-by-number of children fixed effects and observables. Similarly, our identifying assumption for γ is that the macro tax liability $T_{e,s,t}$ is exogenous, conditional on education-by-state and education-by-year fixed effects and observables. Thus, two independent sources of exogenous variation in tax liabilities are needed. For the micro response β, we require variation in tax liabilities across individuals *within* the same labor market. For the macro response γ, we require variation in average tax liabilities *between* labor markets.

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

[44]Note that without income effects, $\frac{\partial \hat{\mathcal{K}}_i}{\partial T_i} = \frac{\partial \hat{\mathcal{K}}_i}{\partial b}$.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 $\frac{\partial \hat{\mathcal{K}}_i}{\partial T_i} = \frac{\partial \hat{\mathcal{K}}_i}{\partial b} = \frac{\partial \hat{\mathcal{K}}_i}{\partial (T_i + b)}$. For our main specification, we will assume no income effects and therefore estimate directly $\frac{\partial \hat{\mathcal{K}}_i}{\partial (T_i + b)}$ thereby using both variation in T_i and b to estimate the parameter of interest with maximum power. We tested whether the condition $\frac{\partial \hat{\mathcal{K}}_i}{\partial T_i} = \frac{\partial \hat{\mathcal{K}}_i}{\partial b}$ holds and found that the difference was very small and statistically insignificant. We therefore only report results under the no income effect assumption.

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

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

III.3 Empirical Results

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

The top panel of Table 2 shows the IV estimates for the micro participation (Column 1) and employment (Column 2) responses to taxes and transfers based on equation (26) above. The results indicate a clear negative and statistically significant participation effect of taxes, consistent with the prior literature. We find that a $1000 increase in taxes leads to a 3.1 percentage point reduction in the participation probability which translates to an elasticity of -0.57.[45] We also see fairly similar micro responses for employment.[46]

Our elasticity estimates are well within the range of elasticities that is reported in the literature.[47] 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,

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

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

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

most of the tax variation in these papers would also have come from across group variation within labor markets.

The macro participation and employment IV estimates are displayed in the second panel of Table 2. These correspond to empirical estimates from a macro-level (education-state-year cells) 2SLS regression of participation and employment rates on market-level tax liabilities, controlling for education-by-state and education-by-year fixed effects and percent black, percent Hispanic, average age, average age-squared, average number of children and their interactions with education and region-specific time trends. Since the number of observations is much smaller and since there is less variation in tax liabilities across labor markets, the coefficients are estimated less precisely. Nevertheless, there is some suggestive evidence that the macro participation and employment responses are smaller than the micro ones. According to Proposition 3, such a finding is consistent with a matching model where the bargaining power is lower than the one prescribed by the Hosios condition.

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

A concern with our macro estimates, which are identified by state-year variation in tax liabilities, is that they may be confounded by policy endogeneity. In particular, states may endogenously set taxes and welfare benefits based on prevailing local economic conditions. Our baseline estimates control for region-specific time trends which should partially address this issue. To further explore the robustness of our estimates, we consider several alternative specifications and report the results in Table 3. Table 3 provides a series of robustness tests. The first column reports our baseline estimates for comparison. In columns 2-4 we drop the region-specific time trends from the regressions and include alternative controls for pre-trends. Since the micro participation regressions control for year-by-state fixed effects, these are not affected (Panel A), but Panel B and C show that the macro responses are very robust to controlling for division-by-year fixed effects, region-by-year fixed effects and no controls for pre-trends. In column 5 we present our results dropping state taxes (state EITC and state income taxes) from our imputed tax liability and instrument, as those may be endogenous, as Hoynes and Patel (2015) have argued. While this slightly reduces the precision of our macro estimates, the results are qualitatively similar. Finally, Column 6 controls for the state unemployment rate interacted with education as a proxy for the state specific economic environment and shows a very similar pattern. Overall, the robustness of our estimates suggest that policy endogeneity is not of first-order importance in our setting.[48]

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

Overall, these results suggest that while micro labor supply responses are sizable and in line with what the literature has found before, they may not always be good approximations for the macro employment responses. In particular our evidence broadly suggests that macro responses tend to be lower than micro responses. Although this is some of the first evidence on the gap between micro and macro elasticities, it is however worth noting that our macro estimates are less precisely estimated than our micro ones. Such discrepancy can easily been explained by the limited policy variations at the state level over time, compared to policy variations across individuals with different number of kids over time. Future research should use other source of policy variations as robustness checks for our macro estimates.

IV Simulating the Optimal Tax Schedule

In this section we show how unemployment and wage responses affect the shape of the optimal tax and transfer schedule. For this purpose we simulate the optimal tax schedule using the sufficient statistics formula for the optimal tax and transfer schedule. In line with the empirical section, we focus on the no-cross effects model with its restricted set of sufficient statistics. These simulations are very stylized and should be viewed as an illustration of the comparative statics of our optimal tax formula, that highlight the importance of taking spillovers into account. The resulting tax schedule should not be viewed as a precise attempt to derive the optimal tax schedule for any particular population.[49]

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

[49] Such an exercise for the U.S. would, for example, have to take into account that policy makers seem to have placed different welfare weights on different groups of single individuals, depending on the number of children. Backing out the implicit welfare weights in the current tax schedule given an optimal tax framework and calibrating how the tax

To simulate the optimal tax schedule, we solve the system of first-order conditions derived in the theoretical 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 (9) and (12) 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 (14).[50] We partition the income distribution into discrete bins, corresponding to the zero income level, the 3 education groups in our empirical analysis, as well as a 4th group: singles with Bachelor degrees, which we did not use in our empirical analysis due to the lack of identifying policy variation for this group. We take the average number of individuals over all years as the population shares of the education groups and assign to each group the average income over our sample period. 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 redistribution. We set $\nu = 0.5$, which leads to optimal tax schedule similar to the observed schedules, but in the Online Appendix we also report results for $\nu = 1$. 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:

$$h_i = h_i^0 + \frac{\partial \mathcal{H}_i}{\partial (T_i + b)} \left((T_i + b) - \left(T_i^0 + b^0 \right) \right). \tag{28}$$

We present simulations of the optimal tax schedule based on the formula derived in this paper, which we refer to as the KKLS formula, and contrast this tax schedule with simulations based on the optimal tax formula in Saez (2002).

Figure 2a) shows the optimal tax and transfer schedule for the lowest 3 education groups using the employment and participation response estimates from our empirical section. The dashed line with circles shows the optimal tax schedule implied by our no-cross effects welfare formula, which relies on 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 solid line with stars 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 (2002) formula with micro employment responses, our formula implies a lump sum transfer to the non-employed about 45 percent larger and significantly higher marginal tax rates (a flatter slope). This is because our estimates imply lower

schedule given these welfare weights would change under alternative models would be very interesting, but beyond the scope of this paper.

[50] 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.

macro than micro participation responses, so that the spillover effects attenuate the welfare gain of a transfer to the working poor. The Saez (2002) formula calibrated with macro employment responses implies larger transfers at the bottom than when micro employment responses are used for calibration and a somewhat flatter slope. This is because we estimate larger micro employment responses than macro ones. To highlight the differences in the slopes, Figure 2b) shows the implied employment tax rates, i.e. $\frac{T_i+b}{w_i}$, at each income level. Clearly the Saez (2002) formula with micro employment effects generates the lowest employment tax rate, which is in fact negative like the EITC. Saez (2002) with macro employment effects, generates larger employment tax rates that is very close to zero for the lowest income group and finally the KKLS optimal tax formula yields an employment tax rate that for the lowest income group is positive, thus resembling more an NIT.

In Figure 3a) we show how, holding the macro employment response constant, the macro-micro participation ratio affects the optimal tax schedule. The line with circles shows the benchmark tax schedule from Figure 2 using our optimal tax formula with our main empirical estimates. The line with stars shows the optimal tax schedule using our formula when we double the macro-micro participation ratio but 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 and the optimal tax is a clearly EITC-like schedule, as Figure 3b) shows the employment tax rate is indeed negative at the bottom. The line with plus signs on the other hand shows the optimal tax schedule when we cut the macro-micro participation ratio to 0.5, thus leading to large negative spillovers where the macro response is smaller than the micro response. This makes the overall tax profile much flatter and the benefits to the non-employed larger, mirroring an NIT situation.

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, especially during recessions. Our estimates in Table 4, while noisy, are consistent with this view: while both macro and micro responses decline in recessions, the decline is much larger for macro responses, both with respect to employment and participation. The business cycle macro estimates suggest that spillover effects could be larger during economic downturns. Figure 4 simulates how the optimal tax schedule would vary over the business cycle given our estimates from Table 4. We present results from the estimates based on the 6 month change in the unemployment rate here, but using the other measures yields qualitatively very similar results. In Figure 4a) and 4b) we show the optimal tax schedule for different business cycle states implied by our (KKLS) optimal tax formula. The transfer at zero income is around 4300 USD during a strong labor market with a negative employment tax of about -10 percent for moving from zero income to the first income group. During weak labor markets the simulation suggests that the transfer at zero should increase to 7600 USD per year with a much higher employment tax of about 30 percent. In contrast, panels (c) and (d) of Figure 4 show the tax schedule implied by the Saez (2002) formula using the macro employment effects estimated over the business cycle.[51] While the decline in macro employment responses during weak labor markets also leads to an increase in transfers at the bottom and a slight increase in employment tax rates, the difference is only around 1000 USD due to the absence of the spillover channel.

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

References

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

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

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

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

Bitler, Marianne, Hilary Hoynes, and Elira Kuka, "Do In-Work Tax Credits Serve as a Safety Net?," 2014, (19785, forthcoming *Journal of Human Resources*).

Blundell, Richard, John C Ham, and Costas Meghir, *Unemployment and female labour supply* number 149, Centre for Economic Policy Research, 1987.

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

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

_ , "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

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

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

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

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

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

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

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

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment," *The Quarterly Journal of Economics*, 2013, *128* (2), 531–580.

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

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

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

_ and James Mirrlees, "Optimal Taxation and Public Production," *American Economic Review*, 1971, *61*, 8–27 also 261–278.

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

Eissa, Nada and Hilary W. Hoynes, "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply," *Tax Policy and the Economy*, 2006, *20*, pp. 73–110.

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

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

Fahri, Emmanuel and Iván Werning, "A Theory of Macroprudential Policies in the Presence of Nominal Rigidities," *Econometrica*, 2016, *85* (5), 1645–1704.

Farber, Henry S and Robert G Valletta, "Do extended unemployment benefits lengthen unemployment spells? Evidence from recent cycles in the US labor market," Technical Report, National Bureau of Economic Research 2013.

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

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

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

Hagedorn, Marcus, Fatih Karahan, Iourii Manovskii, and Kurt Mitman, "Unemployment benefits and unemployment in the great recession: the role of macro effects," Technical Report, National Bureau of Economic Research 2013.

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

Hansen, Claus Thustrup, Lars Haagen Pedersen, and Torsten Sløk, "Ambiguous effects of tax progressivity – theory and Danish evidence," *Labour Economics*, 2000, *7* (3), 335–347.

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

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

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

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

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

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

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

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden, "Optimal income taxation with Kalai wage bargaining and endogenous participation," *Social Choice and Welfare*, 2014, *42* (2), 381–402.

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

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

Keynes, John Maynard, *The General Theory of Employment, Interest and Money*, New York: Macmillan, 1936.

Kroft, Kory, "Takeup, Social Multipliers and Optimal Social Insurance," *Journal of Public Economics*, 2008, *92* (3-4), 623–668.

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

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

Landais, Camille, Pascal Michaillat, and Emmanuel Saez, "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory," *American Economic Journal: Economic Policy*, 2016, p. forthcoming.

Lavecchia, Adam M, "Minimum Wage Policy with Optimal Taxes and Unemployment," *mimeo, University of Toronto*, 2016.

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

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

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

Leigh, Andrew, "Who Benefits from the Earned Income Tax Credit? Incidence among Recipients, Coworkers and Firms," *The B.E. Journal of Economic Analysis & Policy*, 2010, *10* (1), 1–43.

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

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

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

Marinescu, Ioana, "The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board," Technical Report 2014.

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

Meyer, Bruce and D Sullivan, "The effects of Welfare and Tax Reforms: The Material Well-Being of Single Mothers in the 1980s and 1990s," *Journal of Public Economics*, 2004, *88*, 1387–1420.

__ **and** __ , "Changes in the Consumption Income and well-being of single mothers Headed families," *American Economic Review*, 2008, *98*, 2221–2241.

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

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

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

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

__ **and** __ , "Optimal Public Expenditure with Inefficient Unemployment," 2017, (21322).

Mirrlees, James, "Optimal Marginal Tax Rates at Low Income," *mimeo*, 1999.

Moen, Espen R, "Competitive Search Equilibrium," *Journal of Political Economy*, 1997, *105* (2), 385–411.

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

Mogstad, Magne and Chiara Pronzato, "Are Lone Mothers Responsive to Policy Changes? Evidence from a Workfare Reform in a Generous Welfare State," *The Scandinavian Journal of Economics*, 2012, *114* (4), 1129–1159.

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

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

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

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

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

__ , *Equilibrium Unemployment Theory, 2nd Edition*, Vol. 1 of MIT Press Books, The MIT Press, 2000.

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

Ramsey, Frank P., "A Contribution to the Theory of Taxation," *The Economic Journal*, 1927, *37* (145), 47–61.

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

— , "Unemployment insurance and job search in the Great Recession," Technical Report, National Bureau of Economic Research 2011.

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

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

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

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

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

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

A Theoretical Appendix

A.1 Derivation of Equation (12)

Differentiating (7) 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 $\mathscr{U}_i(\mathbf{t}) \equiv \mathscr{P}_i(\mathbf{t}) \; (u(\mathscr{C}_i(\mathbf{t})) - d_i) + (1 - \mathscr{P}_i(\mathbf{t}))\, u(b)$ with respect to b gives:

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

Using $h_0 = 1 - \sum_{i=1}^{I} h_i$ and Equation (13) leads to (12).

Using Equations (3) and (4), Equation (9) can be rewritten as:

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

Using that $\frac{\partial \mathscr{C}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial T_i} - 1$ and for $j \neq i$, $\frac{\partial \mathscr{C}_i}{\partial T_i} = \frac{\partial \mathscr{W}_i}{\partial T_i}$, summing the latter equation for all $j \in \{1, ..., I\}$ and subtracting this sum by Equation (12) leads to:

$$\begin{aligned} 0 \quad = \quad & \sum_{i=0}^{I} h_i + \sum_{i=1}^{I} (T_i + b) \left(\sum_{j=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_j} - \frac{\partial \mathcal{H}_i}{\partial b} \right) - \left(g_0 h_0 + \sum_{i=1}^{I} g_i h_i \right) \\ + \quad & \sum_{i=1}^{I} g_i h_i \left(\sum_{j=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} - \frac{\partial \mathscr{W}_i}{\partial b} \right) + \sum_{i=1}^{I} g_i h_i \frac{u(c_i) - d_i - u(b)}{u'(c_i)} \left(\sum_{j=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_j} - \frac{\partial \mathscr{P}_i}{\partial b} \right) \end{aligned} \qquad (29)$$

In the absence of income effects, a simultaneous change in all tax liabilities and welfare benefit $\Delta T_1 = ... = \Delta T_i = -\Delta b$ induces no changes in wages, conditional employment probabilities not employment levels, so that $\sum_{j=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} = \frac{\partial \mathscr{W}_i}{\partial b}$, $\sum_{j=1}^{I} \frac{\partial \mathscr{P}_i}{\partial T_j} = \frac{\partial \mathscr{P}_i}{\partial b}$ and $\sum_{j=1}^{I} \frac{\partial \mathcal{H}_i}{\partial T_j} = \frac{\partial \mathcal{H}_i}{\partial b}$. Plugging these equalities in (29) leads to: $g_0 h_0 + \sum_{i=1}^{I} g_i h_i = \sum_{i=0}^{I} h_i = 1$.

A.2 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 $q_i = Q_i(\theta_i) \stackrel{\text{def}}{\equiv} \mathcal{M}_i(\nu_i, k_i)/\nu_i = \mathcal{M}_i(1, 1/\theta_i)$, which is decreasing in tightness $\theta_i \stackrel{\text{def}}{\equiv} \nu_i/k_i$. 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 $p_i = P(\theta_i) \stackrel{\text{def}}{\equiv} \mathcal{M}_i(\nu_i, k_i)/k_i = \mathcal{M}_i(\theta_i, 1)$. Therefore, the conditional probability p_i is a decreasing function of the gross wage through $p_i = P_i\left(Q_i^{-1}\left(\kappa_i/(y_i - w_i)\right)\right)$, 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 (4). Moreover, we get from $p_i = \mathscr{L}_i(w_i)$ and (4) 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: $\frac{w_i}{p_i}\frac{\partial \mathscr{P}_i}{\partial w_i} = -\frac{1-\mu_i}{\mu_i}\frac{w_i}{y_i - w_i}$. Moreover, when $\beta_i < 1$, Equation (15) implies 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 thus finally get:

$$\frac{\partial \mathscr{U}_i}{\partial T_i} = \left[-1 + (1 - \beta_i)\left(1 - \frac{1-\mu_i}{\mu_i}\frac{\beta_i}{1-\beta_i}\right)\right] p_i = \frac{\beta_i}{\mu_i}\left.\frac{\partial \mathscr{U}_i}{\partial T_i}\right|^{\text{Micro}} \tag{30}$$

A.3 Proof of Proposition 4

Using Equations (3), (4) and (5), the assumption $\frac{\mathbf{d}\mathscr{P}}{\mathbf{dT}} = \mathbf{0}$ implies that $\frac{\mathbf{d}\mathcal{K}}{\mathbf{dT}} \cdot \left(\left.\frac{d\mathcal{K}}{dT}\right|^{\text{Micro}}\right)^{-1} = -\frac{d\mathscr{C}}{dT} = \frac{\mathbf{d}\mathcal{H}}{\mathbf{dT}} \cdot \left(\left.\frac{d\mathcal{H}}{dT}\right|^{\text{Micro}}\right)^{-1}$. Therefore, Equation (10) is equivalent to:

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

As wages equals the marginal product of labor, and the production function exhibits constant returns to scale, one has $\sum_{i=1}^{I} \frac{\partial \mathscr{W}_i}{\partial T_j} h_i = 0$ for all $j \in \{1, ..., I\}$. This means that $\mathbf{h}$ is an eigenvector of matrix $\frac{\mathbf{d}\mathscr{W}}{\mathbf{dT}}$ associated to the eigenvalue 0, thereby an eigenvector of matrix $\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}$ associated to the eigenvalue -1. One therefore get that $\left(\frac{\mathbf{d}\mathscr{C}}{\mathbf{dT}}\right)^{-\mathbf{1}} \cdot \mathbf{h} = -\mathbf{h}$, which ends the proof.

A.4 Optimal Unemployment and Welfare Benefits

This section considers the case of different unemployment and welfare benefits. Let g_i^b denote the welfare weights on the unemployed and let it be defined from the microeconomic effect of a rise in unemployment benefits. We thus get:

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

The first-order condition with respect to b_j writes:

$$0 = -(k_j - h_j) \sum_{i=1}^{I} (T_i + b_i) \frac{\partial \mathcal{H}_i}{\partial b_j} + \sum_{i=1}^{I} (z - b_i) \frac{\partial \mathcal{K}_i}{\partial b_j} + \sum_{i=1}^{I} \frac{\partial \mathscr{U}_i}{\partial b_j} \left(\left.\frac{\partial \mathscr{U}_i}{\partial b_i}\right|^{\text{Micro}} \right)^{-1} g_i^b (k_i - h_i)$$

The optimal unemployment benefits formula in matrix term is therefore very similar to the corresponding optimal tax formula (18):

$$\mathbf{0} = -(\mathbf{k} - \mathbf{h}) + \frac{\mathbf{d}\mathcal{H}}{\mathbf{db}} \cdot (\mathbf{T} + \mathbf{b}) + \frac{\mathbf{d}\mathcal{K}}{\mathbf{db}} \cdot (\mathbf{z} - \mathbf{b}) - \frac{\mathbf{d}\mathcal{K}}{\mathbf{db}} \cdot \left(\left.\frac{\mathbf{d}\mathcal{K}}{\mathbf{db}}\right|^{\text{Micro}} \right)^{\mathbf{-1}} \cdot \left(\mathbf{g^b}\ (\mathbf{k} - \mathbf{h})\right)$$

Finally, the first-order condition on the welfare benefit z is simply:

$$0 = -k_0 + \sum_{i=1}^{I} (T_i + b_i) \frac{\partial \mathcal{H}_i}{\partial z} + \sum_{i=1}^{I} (z - b_i) \frac{\partial \mathcal{K}_i}{\partial z} + \frac{\partial \Omega}{\partial u(z)} \frac{u'(z)}{\lambda}$$

A.5 Endogenous Search Intensity

Denoting $e_i = \mathscr{E}_i(\mathbf{t})$ the effort choice made, the first-order condition for optimal search is:[52]

$$D'(e_i) = a_i \left[u(c_i) - d_i - u(b) \right] \qquad \Leftrightarrow \qquad e_i\, D'(e_i) - D(e_i) = U_i - u(b)$$

which eventually implies:[53]

$$\mathscr{E}_i(\mathbf{t})\, D'\left(\mathscr{E}_i(\mathbf{t})\right) - D\left(\mathscr{E}_i(\mathbf{t})\right) = \mathscr{U}_i(\mathbf{t}) - u(b) \tag{31a}$$

Using the envelope theorem, we get:

$$\frac{\partial \mathscr{U}_i}{\partial T_j}^{\text{Micro}} = \left.\frac{\partial \mathscr{U}_i}{\partial T_j}\right|_{\mathbf{w,a}} = -e_i\, a_i\, u'(c_i)\, \mathbb{1}_{i=j} \tag{31b}$$

$$\frac{\partial \mathscr{U}_i}{\partial T_j} = e_i\, a_i\, u'(c_i) \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{A}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{a_i\, u'(c_i)} \right] \tag{31c}$$

$$= \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{A}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{a_i\, u'(c_i)} \right] \frac{\partial \mathscr{U}_i}{\partial T_i}^{\text{Micro}}$$

We thus redefine matrix $\mathcal{A}$ as the one whose term in row j and column i is $\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{A}_i}{\partial T_j} \frac{u(c_i) - d_i - u(b)}{a_i\, u'(c_i)}$, we get in matrix terms:

$$\frac{\mathbf{dU}}{\mathbf{dT}} = -\mathcal{A} \cdot \frac{\mathbf{dU}}{\mathbf{dT}}^{\text{Micro}}$$

[52] We here used that: $U_i - u(b) + D(e) = e_i a_i \left[u(c_i) - d_i - u(b) \right]$

[53] The derivative of $e \mapsto e\, D'(e) - D(e)$ is $eD''(e)$ which is positive from the convexity of $D(\cdot)$. Therefore, Equation (31a) uniquely determines the search intensity $\mathscr{E}_i(\mathbf{t})$ in labor market i.

From Equation (31a), search intensity in labor market i is a function denoted $\mathcal{E}_i$ of gross expected utility: $e_i = \mathcal{E}_i(U_i)$, so that $\mathscr{E}_i(\mathbf{t}) \stackrel{\text{def}}{\equiv} \mathcal{E}_i\left(\mathscr{U}_i(\mathbf{t})\right)$. We have that:

$$\begin{aligned}
\frac{\partial \mathscr{E}_i}{\partial T_j}^{\text{Micro}} &= -e_i\, a_i\, u'(c_i)\, \frac{\partial \mathcal{E}_i}{\partial U_i}\, \mathbb{1}_{i=j} \\
\frac{\partial \mathscr{E}_i}{\partial T_j} &= \frac{\partial \mathscr{U}_i}{\partial T_j}\frac{\partial \mathcal{E}_i}{\partial U_i} = \left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{A}_i}{\partial T_j}\frac{u(c_i) - d_i - u(b)}{a_i\, u'(c_i)}\right] e_i\, f_i\, u'(c_i)\frac{\partial \mathcal{E}_i}{\partial U_i} \\
&= -\left[\frac{\partial \mathscr{C}_i}{\partial T_j} + \frac{\partial \mathscr{A}_i}{\partial T_j}\frac{u(c_i) - d_i - u(b)}{a_i\, u'(c_i)}\right]\frac{\partial \mathscr{E}_i}{\partial T_i}^{\text{Micro}}
\end{aligned}$$

so we get in matrix terms:

$$\frac{\mathbf{d}\mathscr{E}}{\mathbf{dT}} = -\mathcal{A}\cdot\frac{\mathbf{d}\mathscr{E}}{\mathbf{dT}}^{\text{Micro}}$$

Therefore, using (5), we have that:

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

which implies Equation (23) from (22).

A.6 Retrieving Landais et al. (2016)

Step 1: Specifying the version of the present model consistent with Landais et al. (2016)

To retrieve the optimal policy rule of Landais et al. (2016), we assume in this Appendix that $I = 1$, there are full participation $k = 1$the social welfare objective to be unweighted utilitarian, so Equation (8) simplifies to $-\left(\left.\frac{\partial \mathscr{U}}{\partial T}\right|^{\text{micro}}\right)^{-1} gh = \frac{1}{\lambda}$ and profits are fully taxed so that $\tau = 1$. As job search intensity is endogenous, individuals solve:

$$\mathscr{U}(T, b) = \max_{e} \qquad e\mathscr{A}(T, b)u\left(\mathscr{W}(T, b) - T\right) + \left(1 - e\mathscr{A}(T, b)\right)u(b) - D(e) \tag{33}$$

Considering the derivatives of $\mathscr{W}$ and $\mathscr{A}$ are nil, we get:

$$\left.\frac{\partial \mathscr{U}}{\partial T}\right|^{\text{Micro}} = -h\, u'(c) \qquad \left.\frac{\partial \mathscr{U}}{\partial b}\right|^{\text{Micro}} - u'(b) = -h\, u'(b) \tag{34}$$

The first-order condition associated to (33) is:

$$D'(e) = \mathscr{A}(T, b)\left[u(w - T) - u(b)\right]$$

which leads to

$$\mathscr{E}(T, b)D'(\mathscr{E}(T, b)) - D(\mathscr{E}(T, b)) = \mathscr{U}(T, b) - u(b)$$

and so:

$$\frac{\dfrac{\partial \mathscr{E}(T,b)}{\partial T}}{\left.\dfrac{\partial \mathscr{E}(T,b)}{\partial T}\right|^{\text{Micro}}} = \frac{\dfrac{\partial \mathscr{U}(T,b)}{\partial T}}{\left.\dfrac{\partial \mathscr{U}(T,b)}{\partial T}\right|^{\text{Micro}}} \quad , \quad \frac{\dfrac{\partial \mathscr{E}(T,b)}{\partial b}}{\left.\dfrac{\partial \mathscr{E}(T,b)}{\partial b}\right|^{\text{Micro}}} = \frac{\dfrac{\partial \mathscr{U}(T,b)}{\partial b} - u'(b)}{\left.\dfrac{\partial \mathscr{U}(T,b)}{\partial b}\right|^{\text{Micro}} - u'(b)}$$

Combing the latter equations with (34) leads to:

$$\frac{\partial \mathscr{U}}{\partial T} = -h\,u'(c)\frac{\dfrac{\partial \mathscr{E}(T,b)}{\partial T}}{\left.\dfrac{\partial \mathscr{E}(T,b)}{\partial T}\right|^{\text{Micro}}} \quad , \quad \frac{\partial \mathscr{U}}{\partial b} = -h\,u'(b)\frac{\dfrac{\partial \mathscr{E}(T,b)}{\partial b}}{\left.\dfrac{\partial \mathscr{E}(T,b)}{\partial b}\right|^{\text{Micro}}} + u'(b) \tag{35}$$

The government's Lagrangian is:

$$\Lambda(T,b) \stackrel{\text{def}}{\equiv} (T+b)\,\mathcal{H}(T,b) - b - E + \Pi(T,b) + \frac{1}{\lambda}\mathscr{U}(T,b)$$

where

$$\Pi(T,b) \stackrel{\text{def}}{\equiv} \max_{h} \qquad F(h) - \left(\mathscr{W}(T,b) + J\left(\mathscr{A}(T,b)\right)\right)h$$

with first-order and envelope conditions:

$$F'(h) = w + J \qquad \frac{\partial \Pi}{\partial T} = -\left(\frac{\partial \mathscr{W}}{\partial T} + J'(a)\frac{\partial \mathscr{A}}{\partial T}\right)h \qquad \frac{\partial \Pi}{\partial b} = -\left(\frac{\partial \mathscr{W}}{\partial b} + J'(a)\frac{\partial \mathscr{A}}{\partial b}\right)h$$

The first-order conditions associated to the government's program are:[54]

$$T \quad : \qquad 0 = h + (T+b)\frac{\partial \mathcal{H}}{\partial T} - \left(\frac{\partial \mathscr{W}}{\partial T} + J'(a)\frac{\partial \mathscr{A}}{\partial T}\right)h + \frac{1}{\lambda}\frac{\partial \mathscr{U}}{\partial T} \tag{36a}$$

$$b \quad : \qquad 0 = h - 1 + (T+b)\frac{\partial \mathcal{H}}{\partial b} - \left(\frac{\partial \mathscr{W}}{\partial b} + J'(a)\frac{\partial \mathscr{A}}{\partial b}\right)h + \frac{1}{\lambda}\frac{\partial \mathscr{U}}{\partial b} \tag{36b}$$

Using (35) and $\frac{\partial \mathscr{C}}{\partial T} = \frac{\partial \mathscr{W}}{\partial T} - 1$, the system (36) can be rewritten as:

$$T: \quad 0 \quad = \quad (T+b)\frac{\partial \mathcal{H}}{\partial T} - \frac{\partial \mathscr{C}}{\partial T}h - J'(a)\frac{\partial \mathscr{A}}{\partial T}h - \frac{u'(c)}{\lambda}\frac{\dfrac{\partial \mathscr{E}}{\partial T}}{\left.\dfrac{\partial \mathscr{E}}{\partial T}\right|^{\text{Micro}}}h \tag{37a}$$

$$b: \quad 0 \quad = \quad -(1-h) + (T+b)\frac{\partial \mathcal{H}}{\partial b} - \left(\frac{\partial \mathscr{W}}{\partial b} + J'(a)\frac{\partial \mathscr{A}}{\partial b}\right)h + \frac{u'(b)}{\lambda}\left(1 - h\frac{\dfrac{\partial \mathscr{E}}{\partial b}}{\left.\dfrac{\partial \mathscr{E}}{\partial b}\right|^{\text{Micro}}}\right) \tag{37b}$$

Step 2: Relating behavioral responses with respect to T and b to the responses in Landais et al. (2016) in terms of $\Delta = u(c) - u(b)$

Landais et al. (2016) assume that T and b affect the wage only through the difference Δ of current utility:

$$\Delta(T,b) \stackrel{\text{def}}{\equiv} u\left(\mathscr{C}(T,b)\right) - u(b) = u\left(\mathscr{W}(T,b) - T\right) - u(b) \tag{38}$$

[54]Equation (36a) can be directly retrieved from Equation (21) as $I = 1$, $\tau = 1$ and $\Omega \equiv U$ so that $-\left(\left.\frac{\partial \mathscr{U}}{\partial T}\right|^{\text{micro}}\right)^{-1} gh = \frac{1}{\lambda}$. Equation (36b) follows the extension of (12) for the inclusion of profits.

We now try to rewrite the behavioral responses in (37) in terms of responses to Δ. We denote the latter with a hat. Differentiating (38) implies:

$$d\Delta = u'(c)\left(\frac{\partial \mathscr{W}}{\partial T}-1\right)dT+\left(u'(c)\,\frac{\partial \mathscr{W}}{\partial b}-u'(b)\right)db \tag{39}$$

Let $\hat{\mathscr{W}}(\cdot)$ be the function expressing the gross wage as a function of Δ, we get for any dT and db:

$$dw=\frac{\partial \mathscr{W}}{\partial T}\,dT+\frac{\partial \mathscr{W}}{\partial b}db=\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,d\Delta=\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\left[u'(c)\left(\frac{\partial \mathscr{W}}{\partial T}-1\right)dT+\left(u'(c)\frac{\partial \mathscr{W}}{\partial b}-u'(b)\right)db\right]$$

As these equalities have to hold for each dT and db, we get:[55]

$$\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)\left(\frac{\partial \mathscr{W}}{\partial T}-1\right)=\frac{\partial \mathscr{W}}{\partial T} \quad \Rightarrow \quad \frac{\partial \mathscr{W}}{\partial T}=\frac{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1} \qquad \frac{\partial \mathscr{C}}{\partial T}=\frac{1}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1} \tag{40a}$$

$$\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\left[u'(c)\,\frac{\partial \mathscr{W}}{\partial b}-u'(b)\right]=\frac{\partial \mathscr{W}}{\partial b} \quad \Rightarrow \quad \frac{\partial \mathscr{W}}{\partial b}=\frac{\partial \mathscr{C}}{\partial b}=\frac{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1} \tag{40b}$$

We get from (39):

$$\begin{aligned} d\Delta &= u'(c)\left(\frac{\partial \mathscr{W}}{\partial T}-1\right)dT+\left(u'(c)\,\frac{\partial \mathscr{W}}{\partial b}-u'(b)\right)db \\ &= \frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}dT+\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}db \end{aligned}$$

Moreover, for $m=a,h,e$, as $\mathscr{M}=\mathscr{A},\mathcal{H},\mathscr{E}$ depends on T and b only through changes in Δ as described by functions $\hat{\mathscr{M}}=\hat{\mathscr{A}},\hat{\mathcal{H}},\hat{\mathscr{E}}$, we get:

$$dm=\frac{\partial \mathscr{M}}{\partial T}dT+\frac{\partial \mathscr{M}}{\partial b}db \quad = \quad \frac{\partial \hat{\mathscr{M}}}{\partial \Delta}d\Delta=\frac{\partial \hat{\mathscr{M}}}{\partial \Delta}\left[\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}dT+\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}db\right]$$

So

$$\frac{\partial \mathcal{H}}{\partial T}=\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} \qquad \frac{\partial \mathcal{H}}{\partial b}=\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} \tag{40c}$$

$$\frac{\partial \mathscr{A}}{\partial T}=\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathscr{A}}}{\partial \Delta} \qquad \frac{\partial \mathscr{A}}{\partial b}=\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathscr{A}}}{\partial \Delta} \tag{40d}$$

$$\frac{\partial \mathscr{E}}{\partial T}=\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} \qquad \frac{\partial \mathscr{E}}{\partial b}=\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1}\,\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} \tag{40e}$$

[55] To get the usual property that $\frac{\partial \mathscr{W}}{\partial T}>0>\frac{\partial \mathscr{C}}{\partial T}$, one needs $\frac{\partial \mathscr{W}}{\partial \Delta}<0$. Hence, we have $\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}\,u'(c)-1<0$.

We repeat this exercise for microeconomic responses. Differentiating (38) implies:

$$d\Delta^{\text{Micro}} = -u'(c)\, dT - u'(b)\, db$$

so

$$\begin{aligned} de^{\text{Micro}} = \left.\frac{\partial \mathscr{E}}{\partial T}\right|^{\text{Micro}} dT + \left.\frac{\partial \mathscr{E}}{\partial b}\right|^{\text{Micro}} db &= \left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}} d\Delta^{\text{Micro}} \\ &= \left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}} \left[-u'(c)\, dT - u'(b)\, db\right] \end{aligned}$$

so

$$\left.\frac{\partial \mathscr{E}}{\partial T}\right|^{\text{Micro}} = -u'(c) \left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}} \qquad \left.\frac{\partial \mathscr{E}}{\partial b}\right|^{\text{Micro}} = -u'(b) \left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}} \tag{40f}$$

Step 3: Rewriting our optimal tax formulas Step 1 using behavioral responses of Step 2

We now rewrite the equations in the system (37) using (40). For Equation (37a), we successively get:

$$\begin{aligned} 0 &= (T+b)\frac{\partial \mathcal{H}}{\partial T} - \frac{\partial \mathscr{C}}{\partial T}h - J'(a)\frac{\partial \mathscr{A}}{\partial T}h - \frac{u'(c)}{\lambda}\frac{\frac{\partial \mathscr{E}}{\partial T}}{\left.\frac{\partial \mathscr{E}}{\partial T}\right|^{\text{Micro}}}h \\ &= (T+b)\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} - \frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{h}{u'(c)} - J'(a)\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathcal{A}}}{\partial \Delta}h - \frac{u'(c)}{\lambda}\frac{\frac{u'(c)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{-u'(c)\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}}}h \\ &= (T+b)\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} - \frac{h}{u'(c)} - J'(a)\frac{\partial \hat{\mathcal{A}}}{\partial \Delta}h + \frac{1}{\lambda}\frac{\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}}}h \end{aligned}$$

which leads to (41a) below. For Equation (37b), we successively get:

$$\begin{aligned} 0 &= -(1-h) + (T+b)\frac{\partial \mathcal{H}}{\partial b} - \left(\frac{\partial \mathscr{W}}{\partial b} + J'(a)\frac{\partial \mathscr{A}}{\partial b}\right)h + \frac{u'(b)}{\lambda}\left(1 - h\frac{\frac{\partial \mathscr{E}}{\partial b}}{\left.\frac{\partial \mathscr{E}}{\partial b}\right|^{\text{Micro}}}\right) \\ 1-h &= (T+b)\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} - \frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}h - J'(a)\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathscr{A}}}{\partial \Delta}h \\ &+ \frac{u'(b)}{\lambda}\left(1 - h\frac{\frac{u'(b)}{\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c)-1}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{-u'(b)\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}}}\right) \end{aligned}$$

$$\left(\frac{1-h}{u'(b)} - \frac{1}{\lambda}\right)\left(\frac{\partial \hat{\mathscr{W}}}{\partial \Delta}u'(c) - 1\right) = (T+b)\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} - \frac{\partial \hat{\mathscr{W}}}{\partial \Delta}h - J'(a)\frac{\partial \hat{\mathscr{A}}}{\partial \Delta}h + \frac{1}{\lambda}\frac{\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\text{Micro}}}h$$

which leads to (41b)

$$(T+b)\frac{\partial\hat{\mathcal{H}}}{\partial\Delta} - J'(a)\frac{\partial\hat{\mathscr{A}}}{\partial\Delta}h + \frac{1}{\lambda}\frac{\frac{\partial\mathscr{E}}{\partial\Delta}}{\left.\frac{\partial\mathscr{E}}{\partial\Delta}\right|^{\text{Micro}}}h = \frac{h}{u'(c)} \tag{41a}$$

$$= \left(\frac{1-h}{u'(b)} - \frac{1}{\lambda}\right)\left(\frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c) - 1\right) + \frac{\partial\hat{\mathscr{W}}}{\partial\Delta}h \tag{41b}$$

These two equations are mutually consistent only if

$$\begin{aligned}\frac{h}{u'(c)} &= \left(\frac{1-h}{u'(b)} - \frac{1}{\lambda}\right)\left(\frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c) - 1\right) + \frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c)\frac{h}{u'(c)} \\ 0 &= \left(\frac{h}{u'(c)} + \frac{1-h}{u'(b)} - \frac{1}{\lambda}\right)\left(\frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c) - 1\right)\end{aligned}$$

As $\frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c) < 0$ and so $\frac{\partial\hat{\mathscr{W}}}{\partial\Delta}u'(c) - 1 < 0$, we must have:

$$\frac{h}{u'(c)} + \frac{1-h}{u'(b)} = \frac{1}{\lambda} \tag{42}$$

which is Equation (12) in Landais et al. (2016), i.e. the inverse Euler equation which is usual in moral hazards models and in New Dynamic Public Finance models whenever utility function are additively separable between consumption and effort, as this is here the case.

Step 4:

Following LMS, we define :

$$R \stackrel{\text{def}}{\equiv} 1 - \frac{c-b}{w} = \frac{T+b}{w} \qquad \varepsilon^{\text{Micro}} \stackrel{\text{def}}{\equiv} \frac{\Delta}{1-h}\left.\frac{\partial\hat{\mathcal{H}}}{\partial\Delta}\right|^{\text{Micro}} \qquad \varepsilon^{\text{Macro}} \stackrel{\text{def}}{\equiv} \frac{\Delta}{1-h}\frac{\partial\hat{\mathcal{H}}}{\partial\Delta} \qquad \varepsilon^{\text{f}} \stackrel{\text{def}}{\equiv} \frac{a}{e}\frac{\partial e}{\partial a} \tag{43}$$

Remark 1. *Rewriting the first-order condition on individuals' job search intensity as $D'(e) = a\ \Delta$, and using $h = a\ e$, Landais et al. (2016) should have $\varepsilon^f = \varepsilon^{Micro}\frac{1-h}{h}$.*

Proof: The first-order condition $D'(e) = a\ \Delta$ on individuals' job search intensity from program (33) implies $\log(D'(e)) = \log(a) + \log(\Delta)$. Differentiating implies:

$$\begin{aligned}\frac{D''(e)\ de}{D'(e)} &= \frac{da}{a} + \frac{d\Delta}{\Delta} \\ \frac{D''(e)\ e}{D'(e)}\frac{de}{e} &= \frac{da}{a} + \frac{d\Delta}{\Delta} \\ \frac{de}{e} &= \frac{D'(e)}{D''(e)\ e}\left(\frac{da}{a} + \frac{d\Delta}{\Delta}\right)\end{aligned}$$

Hence:

$$\frac{D'(e)}{D''(e)\ e} = \frac{\Delta}{e}\left.\frac{\partial\hat{\mathscr{E}}}{\partial\Delta}\right|^{\text{Micro}} = \frac{\Delta}{h}\left.\frac{\partial\hat{\mathcal{H}}}{\partial\Delta}\right|^{\text{Micro}} = \frac{1-h}{h}\varepsilon^{\text{Micro}} = \frac{a}{e}\frac{\partial e}{\partial a} = \varepsilon^{\text{f}}$$

Hence

$$\varepsilon^{\mathrm{f}} = \frac{1-h}{h}\varepsilon^{\mathrm{Micro}} \tag{44}$$

□

Let use denote:

$$\alpha \stackrel{\text{def}}{\equiv} \frac{\Delta}{a}\frac{\partial a}{\partial \Delta}$$

From $h = a\, e$, thereby $\frac{dh}{h} = \frac{da}{a} + \frac{de}{e}$, we get:

$$\frac{\Delta}{h}\left.\frac{\partial \hat{\mathcal{H}}}{\partial \Delta}\right|^{\mathrm{Micro}} = \frac{\Delta}{e}\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\mathrm{Micro}} \qquad \Rightarrow \qquad \frac{\Delta}{e}\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\mathrm{Micro}} = \frac{1-h}{h}\varepsilon^{\mathrm{Micro}}$$

and

$$\frac{\Delta}{h}\frac{\partial \hat{\mathcal{H}}}{\partial \Delta} = \frac{\Delta}{a}\frac{\partial \hat{\mathscr{A}}}{\partial \Delta} + \frac{\Delta}{e}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} \qquad \Rightarrow \qquad \varepsilon^{\mathrm{Macro}} = \frac{h}{1-h}\alpha + \frac{h}{1-h}\frac{\Delta}{e}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}$$

Finally, applying the chain rule, we get:

$$\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} = \frac{\partial \hat{\mathscr{E}}}{\partial a}\frac{\partial \hat{A}}{\partial \Delta} + \left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\mathrm{Micro}} \qquad \Rightarrow \qquad \frac{\Delta}{e}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} = \varepsilon^{\mathrm{f}}\,\alpha + \frac{1-h}{h}\varepsilon^{\mathrm{Micro}}$$

Hence:

$$\begin{aligned}
\varepsilon^{\mathrm{Macro}} &= \frac{h}{1-h}\alpha + \frac{h}{1-h}\frac{\Delta}{e}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta} = \frac{h}{1-h}\alpha + \frac{h}{1-h}\varepsilon^{\mathrm{f}}\,\alpha + \varepsilon^{\mathrm{Micro}} = \frac{h}{1-h}\alpha(1+\varepsilon^{\mathrm{f}}) + \varepsilon^{\mathrm{Micro}} \\
&\Rightarrow \qquad \alpha = \frac{1-h}{h}\frac{\varepsilon^{\mathrm{Macro}} - \varepsilon^{\mathrm{Micro}}}{1+\varepsilon^{\mathrm{f}}} && (45) \\
\frac{\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\mathrm{Micro}}} &= \frac{\frac{h}{1-h}\frac{\Delta}{e}\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}}{\frac{h}{1-h}\frac{\Delta}{e}\left.\frac{\partial \hat{\mathscr{E}}}{\partial \Delta}\right|^{\mathrm{Micro}}} = \frac{\varepsilon^{\mathrm{f}}\frac{h}{1-h}\alpha + \varepsilon^{\mathrm{Micro}}}{\varepsilon^{\mathrm{Micro}}} = 1 + \frac{\varepsilon^{\mathrm{f}}}{1+\varepsilon^{\mathrm{f}}}\left(\frac{\varepsilon^{\mathrm{Macro}}}{\varepsilon^{\mathrm{Micro}}} - 1\right) && (46)
\end{aligned}$$

Therefore, we have to retrieve Equation (23) in Landais et al. (2016) from (41a):

$$
\begin{aligned}
\frac{T+b}{h}\frac{\partial\hat{\mathcal{H}}}{\partial\Delta} &= \frac{1}{u'(c)} - \frac{1}{\lambda}\frac{\frac{\partial\mathscr{E}}{\partial\Delta}}{\left.\frac{\partial\mathscr{E}}{\partial\Delta}\right|^{\text{Micro}}} + J'(a)\frac{\partial\hat{\mathscr{A}}}{\partial\Delta} \\
\frac{T+B}{w}\frac{w}{\Delta}\frac{1-h}{h}\frac{\Delta}{1-h}\frac{\partial\hat{\mathcal{H}}}{\partial\Delta} &= \frac{1}{u'(c)} - \frac{1}{\lambda} + \frac{1}{\lambda}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] + \frac{a\,J'(a)}{\Delta}\frac{\Delta}{a}\frac{\partial\hat{\mathscr{A}}}{\partial\Delta} \\
R\frac{w}{\Delta}\frac{1-h}{h}\varepsilon^{\text{Macro}} &= \frac{1}{u'(c)} - \frac{h}{u'(c)} - \frac{1-h}{u'(b)} + \frac{1}{\lambda}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] + \frac{a\,J'(a)}{\Delta}\frac{1-h}{h}\frac{\varepsilon^{\text{Macro}}-\varepsilon^{\text{Micro}}}{1+\varepsilon^{\text{f}}} \\
R\frac{w}{\Delta}\frac{1-h}{h}\varepsilon^{\text{Macro}} &= (1-h)\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \frac{1}{\lambda}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] + \frac{a\,J'(a)}{\Delta}\frac{1-h}{h}\frac{\varepsilon^{\text{Macro}}-\varepsilon^{\text{Micro}}}{1+\varepsilon^{\text{f}}} \\
R\frac{w}{\Delta}\varepsilon^{\text{Macro}} &= h\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \frac{1}{\lambda}\frac{h}{1-h}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] + \frac{a\,J'(a)}{\Delta}\frac{\varepsilon^{\text{Macro}}-\varepsilon^{\text{Micro}}}{1+\varepsilon^{\text{f}}} \\
R\frac{w}{\Delta}\varepsilon^{\text{Macro}} &= h\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \frac{1}{\lambda}\frac{h}{1-h}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] - \frac{a\,J'(a)}{\Delta}\frac{\varepsilon^{\text{Micro}}}{1+\varepsilon^{\text{f}}}\left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right] \\
R\frac{w}{\Delta}\varepsilon^{\text{Macro}} &= h\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right]\left[\frac{1}{\lambda}\frac{h}{1-h}\frac{\varepsilon^{\text{f}}}{1+\varepsilon^{\text{f}}} - \frac{a\,J'(a)}{\Delta}\frac{\varepsilon^{\text{Micro}}}{1+\varepsilon^{\text{f}}}\right] \\
R\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}} &= \frac{h}{\varepsilon^{\text{Micro}}}\frac{\Delta}{w}\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right]\frac{1}{1+\varepsilon^{\text{f}}}\left[\frac{\Delta}{w\,\lambda}\frac{h}{1-h}\frac{\varepsilon^{\text{f}}}{\varepsilon^{\text{Micro}}} - \frac{a\,J'(a)}{w}\right]
\end{aligned}
$$

Using (44) we get:

$$
\begin{aligned}
R\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}} &= \frac{h}{\varepsilon^{\text{Micro}}}\frac{\Delta}{w}\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right]\frac{1}{1+\varepsilon^{\text{f}}}\left[\frac{\Delta}{w\,\lambda} - \frac{a\,J'(a)}{w}\right] \\
R &= \frac{h}{\varepsilon^{\text{Micro}}}\frac{\Delta}{w}\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right]\frac{1}{1+\varepsilon^{\text{f}}}\left[\frac{\Delta}{w\,\lambda} + (1+\varepsilon^{\text{f}})R - \frac{a\,J'(a)}{w}\right] \quad (47)
\end{aligned}
$$

which differs from Equation (23) of Landais et al. (2016) only by the term $\frac{a\,J'(a)}{w}$ instead of $\frac{\eta}{1-\eta}\tau(\theta)$ in Landais et al. (2016). Both terms corresponds to recruiting costs for the firm. However, in our model (as in most of matching models), recruiting costs are cost per vacancy posted. In Landais et al. (2016), these costs are hiring cost of recruiting workers. So both are conceptually identical although slightly different in the details.

Rewriting Landais et al. (2016) in terms of sufficient stat

Using our Equation (44), Equation (47) becomes:

$$
R = \frac{h}{\varepsilon^{\text{Micro}}}\frac{\Delta}{w}\left[\frac{1}{u'(c)} - \frac{1}{u'(b)}\right] + \left[1-\frac{\varepsilon^{\text{Macro}}}{\varepsilon^{\text{Micro}}}\right]\frac{h}{h+(1-h)\varepsilon^{\text{Micro}}}\left(\frac{\Delta}{w\,\lambda} + \frac{h+(1-h)\varepsilon^{\text{Micro}}}{h}R - \frac{a\,J'(a)}{w}\right)
$$

Table 1: Variable Means for Single Individuals

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

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

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

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

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

Table 3: Alternative Estimates of Participation and Employment Responses

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

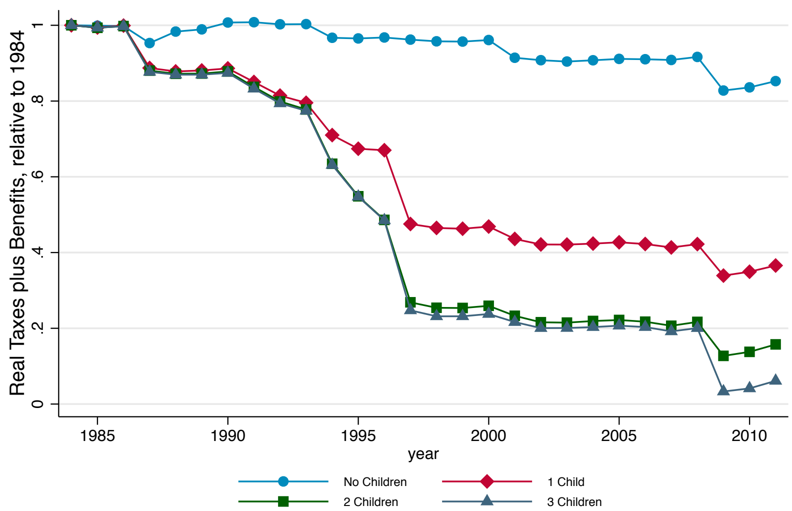
Notes: (* P<.1, ** P<.05, *** P<.01) Standard errors clustered on state level. The sample is restricted to single men and women aged 18-55. The data includes March CPS for 1984-2011 and Outgoing Rotations Groups for 1994-2010. Our baseline specification from Table 3 is contained in column (1). Column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

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

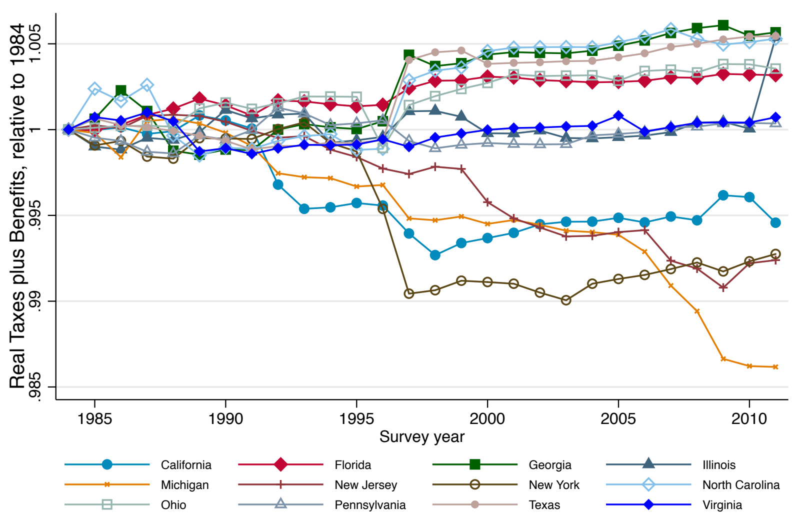
	(1)	(2)	(3)	(4)
	Regression Coef.		**Extrapolated Marg. Effects**	
	Marginal Effect of Tax Liability	Interaction of Tax Liab. with Labor Market Meas.	Weak Labor Market	Strong Labor Market
Panel A: Micro Participation				
6-mo change in unemp	-0.031	0.0011	-0.030	-0.033
	[0.002]***	[0.0004]***		
State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

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

Figure 1: The Variation in Taxes plus Benefits



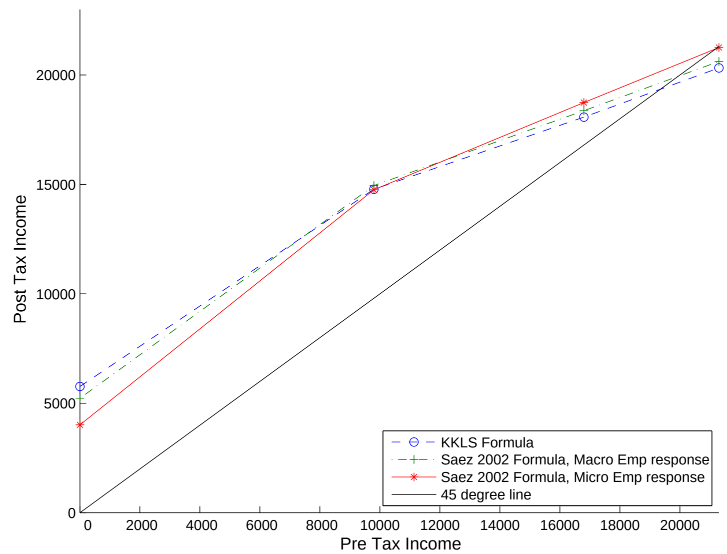
(a) Micro Variation in Taxes plus Benefits



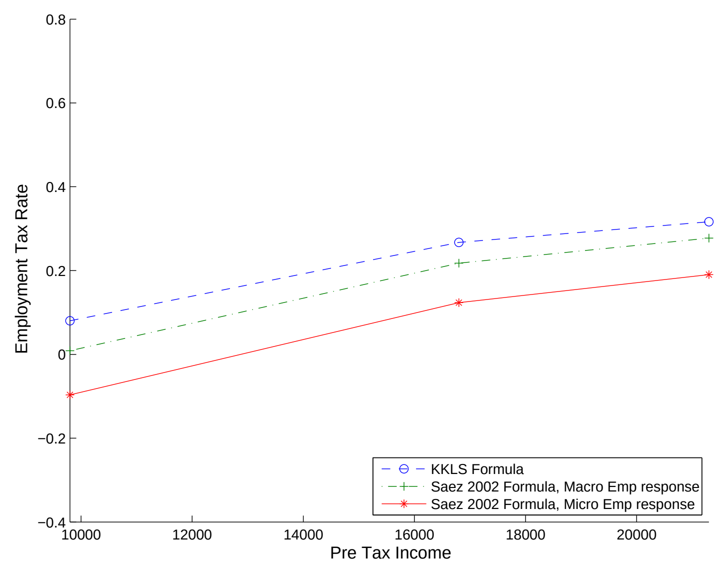
(b) Macro Variation in Taxes plus Benefits

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

Figure 2: Optimal Tax and Transfer Schedule Comparing KKLS Formula with Saez (2002) Formula



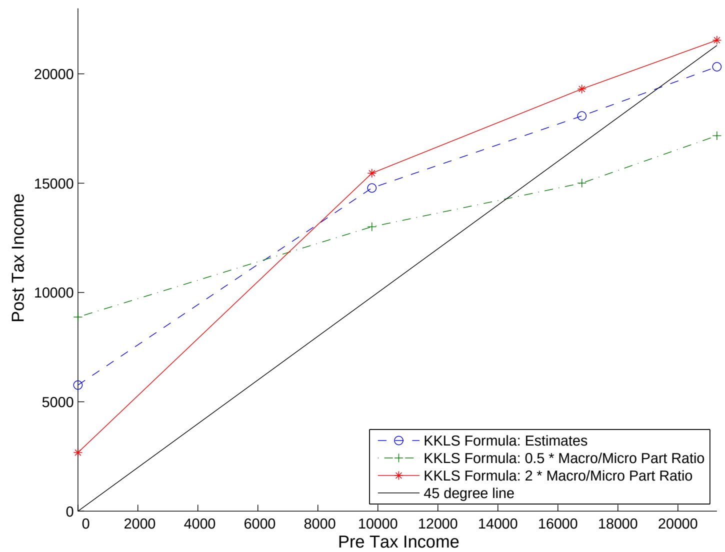
(a) Comparing KKLS vs. Saez (2002) formula: Post vs. Pre-tax income



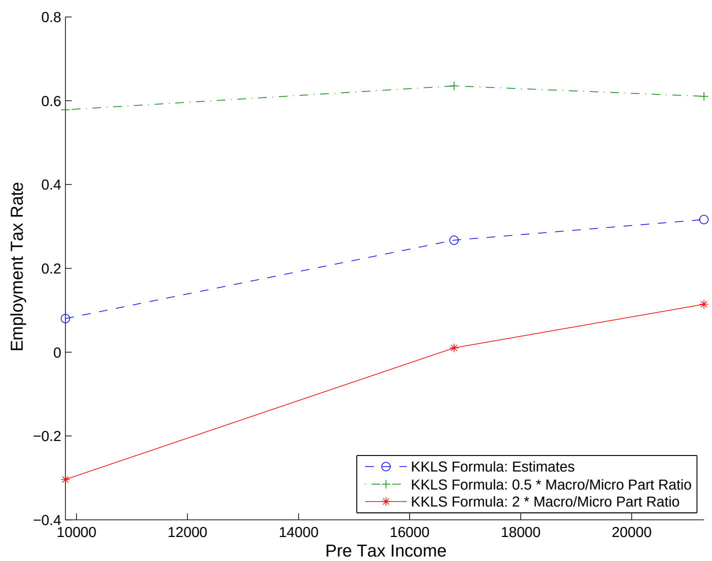
(b) Comparing KKLS vs. Saez (2002) formula: Employment tax rates

Notes: Simulations of the optimal tax and transfer schedule under alternate assumptions on employment and participation responses. The optimal schedule is simulated for 5 income groups, corresponding to the 4 education groups in the empirical section and zero income. Distribution of the income groups is calibrated using CPS data. We show the optimal schedule for the lowest 4 groups where the variation of interest lies. The figure uses the participation and employment responses estimated in the paper. The line with circles uses the optimal welfare formula derived in this paper. The dashed line with plus signs uses the Saez (2002) formula based on the estimated macro responses in this paper, while the solid line uses the estimated micro employment responses in this paper.

Figure 3: The Effect of Changing the Macro Participation Effect on the Optimal Tax and Transfer Schedule



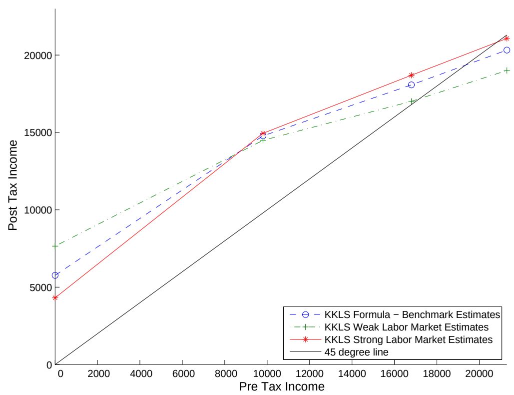
(a) KKLS formula with alternative macro vs micro participation rates: Post vs. Pre-tax income



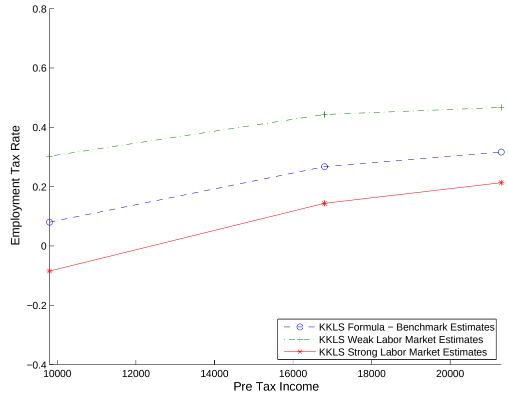
(b) KKLS formula with alternative macro vs micro participation rates: Employment tax rates

Notes: Simulations of the optimal tax and transfer schedule under alternate assumptions on employment and participation responses. The optimal schedule is simulated for 5 income groups, corresponding to the 4 education groups in the empirical section and zero income. Distribution of the income groups is calibrated using CPS data. We show the optimal schedule for the lowest 4 groups where the variation of interest lies. The top figure shows the post vs. pre-tax income relationship while the bottom figure shows the employment tax rates. The line with circles shows the optimal tax schedule given the empirical estimates and the KKLS formula. The solid line shows the optimal schedule if the macro responses are multiplied by 0.5 and the line with plus signs if they are multiplied by 2.

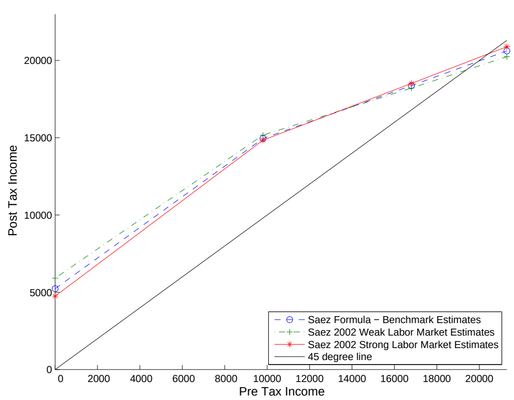
Figure 4: Optimal Tax and Transfer Schedule in Weak vs. Strong Labor Markets



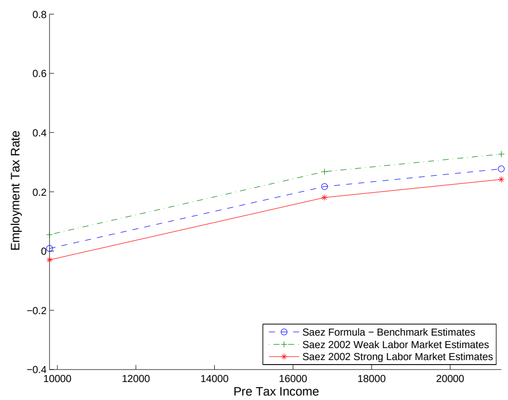
(a) KKLS formula: Post vs. Pre-tax income



(b) KKLS formula: Employment tax rates



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



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

Notes: Simulations of the optimal tax and transfer schedule under alternate macro participation responses. The optimal schedule is simulated for 5 income groups, corresponding to the 4 education groups in the empirical section and zero income. Distribution of the income groups is calibrated using CPS data. We show the optimal schedule for the lowest 4 groups where the variation of interest lies. The top two figures use the KKLS optimal tax formula, the bottom two figures the Saez (2002) optimal tax formula using Macro employment effects. The line with circles corresponds to the benchmark simulation using the estimated, participation and employment responses. The solid line shows the tax schedule using the weak labor market estimates from Table 4 based on the 6 month change in the unemployment rate. The line with plus signs shows the tax schedule for the corresponding strong labor market estimates from Table 4.