Online Appendix for "Using Labor Supply Elasticities To Learn About Income Inequality: The Role of Productivities vs. Preferences"

Katy Bergstrom and William Dodds

A Proofs Appendix

A.1 Proof of Lemma 1

Proof. We prove a slightly stronger statement than stated in the main body (this stronger version will be used in Appendix A.4). We show that if the tax schedule is piece-wise linear (as opposed to linear, as assumed in the main body), then for all (n, α) such that optimal earnings $z^*(n, \alpha)$ is not a kink point of the tax schedule and $z^*(n, \alpha)$ is the unique optimal earnings level for type (n, α), the Jacobian matrix of $G(\log(n), \log(\alpha))$ is given by the following expression:[1]

$$\mathbf{J}_G(\log(n), \log(\alpha)) = \begin{bmatrix} \frac{\partial \log(z^*)}{\partial \log(n)} & \frac{\partial \log(z^*)}{\partial \log(\alpha)} \\ \frac{\partial \log(h^*)}{\partial \log(n)} & \frac{\partial \log(h^*)}{\partial \log(\alpha)} \end{bmatrix} (\log(n), \log(\alpha)) = \begin{bmatrix} 1 + \xi_z^u & \xi_z^c \\ \xi_h^u & \xi_h^c \end{bmatrix} (\log(n), \log(\alpha))$$

This stronger statement implies that if the tax schedule is linear, then the above expression for $\mathbf{J}_G(\log(n), \log(\alpha))$ holds globally. For any individual not locating at a kink point of the piece-wise linear tax schedule, the tax schedule is locally linear with tax rate $(1-T')$ and virtual income R. For any such individual, consider the first order conditions with respect to h and e, evaluated at the optimal levels h^* and e^*:

$$U_h(h^*, e^*; n, \alpha, 1 - T', R) = \alpha u_c(nh^*e^*(1 - T') + R)ne^*(1 - T') - v_h(h^*, e^*) = 0$$
$$U_e(h^*, e^*; n, \alpha, 1 - T', R) = \alpha u_c(nh^*e^*(1 - T') + R)nh^*(1 - T') - v_e(h^*, e^*) = 0$$

Where as before we define $U(h, e; n, \alpha, 1 - T', R)$ as:

$$U(h, e; n, \alpha, 1 - T', R) \equiv \alpha u(nhe(1 - T') + R) - v(h, e)$$

Now, note that n and $1 - T'$ enter the above equations only multiplicatively as $n(1 - T')$; hence, it can be immediately deduced that the elasticities of h and e with respect to n

[1] If type (n, α) has more than one optimal earnings level, which can occur if the tax schedule has decreasing marginal tax rates, $G(\log(n), \log(\alpha))$ is no longer a function, it is a correspondence.

must be the same as with respect to $1-T'$. Differentiating U_h and U_e with respect to n and multiplying by n we get (noting $c^* = nh^*e^*(1-T') + R$):

$$\alpha u_c(c^*)ne^*(1-T') + \alpha u_{cc}(c^*)ne^*(1-T')^2 z^* + U_{hh}(h^*,e^*)\frac{\partial h^*}{\partial n}n + U_{he}(h^*,e^*)\frac{\partial e^*}{\partial n}n = 0$$
$$\alpha u_c(c^*)nh^*(1-T') + \alpha u_{cc}(c^*)nh^*(1-T')^2 z^* + U_{ee}(h^*,e^*)\frac{\partial e^*}{\partial n}n + U_{eh}(h^*,e^*)\frac{\partial h^*}{\partial n}n = 0$$

Differentiating U_h and U_e with respect to $(1-T')$ and multiplying by $(1-T')$, we have:

$$\begin{aligned}(1)\qquad & \alpha u_c(c^*)ne^*(1-T') + \alpha u_{cc}(c^*)ne^*(1-T')^2 z^* + \\ & U_{hh}(h^*,e^*)\frac{\partial h^*}{\partial(1-T')}(1-T') + U_{he}(h^*,e^*)\frac{\partial e^*}{\partial(1-T')}(1-T') = 0\end{aligned}$$

$$\begin{aligned}(2)\qquad & \alpha u_c(c^*)nh^*(1-T') + \alpha u_{cc}(c^*)nh^*(1-T')^2 z^* + \\ & U_{ee}(h^*,e^*)\frac{\partial e^*}{\partial(1-T')}(1-T') + U_{eh}(h^*,e^*)\frac{\partial h^*}{\partial(1-T')}(1-T') = 0\end{aligned}$$

Hence, comparing terms, we must have that:

$$\frac{\partial h^*}{\partial n}n = \frac{\partial h^*}{\partial(1-T')}(1-T')$$

$$\frac{\partial e^*}{\partial n}n = \frac{\partial e^*}{\partial(1-T')}(1-T')$$

Thus, $\xi_h^n = \xi_h^u$. Finally, noting that $\log(z^*) = \log(n) + \log(h^*) + \log(e^*)$, differentiating with respect to n, and substituting in, we have that:

$$\xi_z^n = 1 + \frac{\partial \log(h^*)}{\partial \log(n)} + \frac{\partial \log(e^*)}{\partial \log(n)} = 1 + \frac{\partial \log(h^*)}{\partial \log(1-T')} + \frac{\partial \log(e^*)}{\partial \log(1-T')} = 1 + \xi_z^u$$

The 1 in the above equalities comes from the endowment effect of increasing n.

Lastly, note that α and $1-T'$ enter the first order conditions multiplicatively as $\alpha(1-T')$ *if we hold consumption constant.* Intuitively, the elasticities of hours worked and earnings with respect to α must be the same as the elasticities with respect to $1-T'$, holding consumption constant. In other words, the elasticities of hours worked and earnings with respect to α must be the same as the *compensated* elasticities with respect to $1-T'$. More concretely, by differentiating U_h and U_e with respect to α and multiplying by α:

$$(3)\qquad \alpha u_c(c^*)ne^*(1-T') + U_{hh}(h^*,e^*)\frac{\partial h^*}{\partial \alpha}\alpha + U_{he}(h^*,e^*)\frac{\partial e^*}{\partial \alpha}\alpha = 0$$

$$\alpha u_c(c^*)nh^*(1-T') + U_{ee}(h^*,e^*)\frac{\partial e^*}{\partial \alpha}\alpha + U_{eh}(h^*,e^*)\frac{\partial h^*}{\partial \alpha}\alpha = 0 \tag{4}$$

Differentiating U_h and U_e with respect to R and multiplying by $z(1-T')$ we find:

$$\alpha u_{cc}(c^*)ne^*(1-T')^2 z^* + U_{hh}(h^*,e^*)\frac{\partial h^*}{\partial R}z^*(1-T') + U_{he}(h^*,e^*)\frac{\partial e^*}{\partial R}z^*(1-T') = 0 \tag{5}$$

$$\alpha u_{cc}(c^*)nh^*(1-T')^2 z^* + U_{ee}(h^*,e^*)\frac{\partial e^*}{\partial R}z^*(1-T') + U_{eh}(h^*,e^*)\frac{\partial h^*}{\partial R}z^*(1-T') = 0 \tag{6}$$

Subtracting Equations 5 and 6 from Equations 1 and 2, respectively:

$$\begin{aligned}&\alpha u_c(c^*)ne^*(1-T') + U_{hh}(h^*,e^*)\left(\frac{\partial h^*}{\partial(1-T')} - \frac{\partial h^*}{\partial R}z^*\right)(1-T') + \\ &U_{he}(h^*,e^*)\left(\frac{\partial e^*}{\partial(1-T')} - \frac{\partial e^*}{\partial R}z^*\right)(1-T') = 0\end{aligned} \tag{7}$$

$$\begin{aligned}&\alpha u_c(c^*)nh^*(1-T') + U_{ee}(h^*,e^*)\left(\frac{\partial e^*}{\partial(1-T')} - \frac{\partial e^*}{\partial R}z^*\right)(1-T') + \\ &U_{eh}(h^*,e^*)\left(\frac{\partial h^*}{\partial(1-T')} - \frac{\partial h^*}{\partial R}z^*\right)(1-T') = 0\end{aligned} \tag{8}$$

Hence, comparing terms in Equations 7 and 8 with Equations 3 and 4, we have that:

$$\frac{\partial h^*}{\partial \alpha}\alpha = \left(\frac{\partial h^*}{\partial(1-T')} - \frac{\partial h^*}{\partial R}z^*\right)(1-T')$$

$$\frac{\partial e^*}{\partial \alpha}\alpha = \left(\frac{\partial e^*}{\partial(1-T')} - \frac{\partial e^*}{\partial R}z^*\right)(1-T')$$

Using the definition of the compensated elasticity for $i = e, h$ from Section II we get that $\xi_h^\alpha = \xi_h^c$ and $\xi_e^\alpha = \xi_e^c$. The relationship $\xi_z^\alpha = \xi_z^c$ follows from $\log(z^*) = \log(n) + \log(h^*) + \log(e^*)$, $\xi_h^\alpha = \xi_h^c$, and $\xi_e^\alpha = \xi_e^c$.

□

A.2 Recovering Optimal Effort from Earnings and Hours Worked

First, under the conditions in Proposition 1, we can invert the relationship between (z^*, h^*) and (n, α) so as to write n and α in terms of z^* and h^*. Hence, we can also write e^* as a function of z^* and h^*. We have that $\log(e^*) = \log(z^*) - \log(h^*) -$

$\log(n(\log(z^*), \log(h^*)))$. Taking partial derivatives of $\log(e^*)$ w.r.t. $\log(h^*)$ and $\log(z^*)$:

$$\frac{\partial \log(e^*)}{\partial \log(h^*)} = -1 - \frac{\partial \log(n)}{\partial \log(h^*)} = -1 + \frac{\xi_z^c}{(1+\xi_z^u)\xi_h^c - \xi_h^u \xi_z^c}$$

$$\frac{\partial \log(e^*)}{\partial \log(z^*)} = 1 - \frac{\partial \log(n)}{\partial \log(z^*)} = 1 - \frac{\xi_h^c}{(1+\xi_z^u)\xi_h^c - \xi_h^u \xi_z^c}$$

The equations for $\partial \log(n)/\partial \log(h^*)$ and $\partial \log(n)/\partial \log(z^*)$ come from inverting the derivative matrix of G in Proposition 1. Using $\xi_z^c = \xi_h^c + \xi_e^c$ and the Slutsky equations $\xi_i^u = \xi_i^c + \eta_i$ for $i = z, h, e$, we can rearrange to yield:

(9)
$$\frac{\partial \log(e^*)}{\partial \log(h^*)} = \frac{\xi_e^c + \xi_e^c \eta_h - \xi_h^c \eta_e}{\xi_h^c + \xi_h^c \eta_e - \xi_e^c \eta_h}$$

(10)
$$\frac{\partial \log(e^*)}{\partial \log(z^*)} = \frac{\xi_z^c \eta_e - \xi_e^c \eta_z}{\xi_z^c(1-\eta_e) - \xi_e^c(1-\eta_z)}$$

Intuitively, the relationship between optimal hours and optimal effort depends on the ratio between ξ_e^c and ξ_h^c. However, there are also terms in Equations 9 and 10 that depend on the income effect parameters. Starting with Equation 9, as optimal hours increases, this creates an income effect (which is negative) that causes optimal effort to fall, hence the term $\xi_e^c \eta_h$ in the numerator. Similarly, as optimal effort rises, this causes an income effect, leading optimal hours to fall, which in turn affects optimal effort, hence the term $-\xi_h^c \eta_e$ in the numerator. Likewise, in the denominator of Equation 9, the converse logic holds, hence the terms $\xi_h^c \eta_e$ and $-\xi_e^c \eta_h$. Turning to Equation 10, increasing optimal earnings leads to an income effect on effort, but increasing effort leads to a further income effect, which decreases optimal effort, hence the terms $\xi_z^c \eta_e$ and $-\xi_e^c \eta_z$. In the denominator of Equation 10, changes in earnings lead to income effects, which further effect earnings via income effects on earnings and income effects on effort.

Finally, if income effects are 0 so that $\xi_i^u = \xi_i^c$, then the above equations simplify to:

$$\frac{\partial \log(e^*)}{\partial \log(h^*)} = \frac{\xi_e^c}{\xi_h^c} = \frac{\xi_z^c - \xi_h^c}{\xi_h^c}$$

$$\frac{\partial \log(e^*)}{\partial \log(z^*)} = 0$$

A.3 Proof of Proposition 2 with A Linear Tax Schedule

Proof. Recall that $G : \mathcal{N} \times \mathcal{A} \to \mathcal{Z}^* \times \mathcal{H}^*$ is the continuously differentiable function that maps each $(\log(n), \log(\alpha))$ to a $(\log(z^*), \log(h^*))$.[2] Our goal is to find the inverse function, $G^{-1} : \mathcal{Z}^* \times \mathcal{H}^* \to \mathcal{N} \times \mathcal{A}$. By Lemma 1, we can recover the Jacobian derivative matrix of the function G, denoted $\mathbf{J}_G$:[3]

$$\mathbf{J}_G(\log(n), \log(\alpha)) = \begin{bmatrix} \frac{\partial \log(z^*)}{\partial \log(n)} & \frac{\partial \log(z^*)}{\partial \log(\alpha)} \\ \frac{\partial \log(h^*)}{\partial \log(n)} & \frac{\partial \log(h^*)}{\partial \log(\alpha)} \end{bmatrix} (\log(n), \log(\alpha)) = \begin{bmatrix} 1 + \xi_z^u & \xi_z^c \\ \xi_h^u & \xi_h^c \end{bmatrix} (\log(n), \log(\alpha))$$

We want to show that the mapping G is bijective, i.e., that each (n, α) chooses a unique optimal (z^*, h^*) and each (z^*, h^*) is chosen by a unique (n, α). In order to show that G is bijective, we need to first show that its Jacobian has an everywhere non-zero determinant, which is necessary for local invertibility. We showed in the proof to Proposition 1 that, under the conditions stated in the proposition, $\mathbf{J}_G$ has an everywhere positive determinant. Moreover, $(1 + \xi_z^u) > 0$ (as $1 + \xi_z^u - \xi_z^c > 0$ and $\xi_z^c > 0$) and $\xi_h^c > 0$ so that $\mathbf{J}_G$ has all positive principal minors. Hence $\mathbf{J}_G$ is a P-matrix. A mapping G on a closed rectangular domain characterized by a P-matrix Jacobian must be bijective by Gale and Nikaido (1965) Theorem 4 (we assume the elasticity conditions hold for all $(n, \alpha) \in \mathbb{R}_+^2$ so that the domain is closed and rectangular).

Thus, the mapping G is globally invertible; moreover, by the inverse function theorem, the Jacobian of the inverse mapping G^{-1} is given by:

$$\mathbf{J}_{G^{-1}}(\log(z^*), \log(h^*)) = \begin{bmatrix} \frac{\partial \log(n)}{\partial \log(z^*)} & \frac{\partial \log(n)}{\partial \log(h^*)} \\ \frac{\partial \log(\alpha)}{\partial \log(z^*)} & \frac{\partial \log(\alpha)}{\partial \log(h^*)} \end{bmatrix} (\log(z^*), \log(h^*)) = \begin{bmatrix} 1 + \xi_z^u & \xi_z^c \\ \xi_h^u & \xi_h^c \end{bmatrix}^{-1} (\log(z^*), \log(h^*))$$

From here, we simply pick a particular (z_0^*, h_0^*) and normalize $(\log(n(z_0^*, h_0^*)), \log(\alpha(z_0^*, h_0^*))) = (0, 0)$. Finally, if γ represents a path from $(\log(z_0^*), \log(h_0^*))$ to $(\log(z^*), \log(h^*))$, we have by Stokes' Theorem:[4]

$$\begin{bmatrix} \log(n(z^*, h^*)) \\ \log(\alpha(z^*, h^*)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \int_\gamma \mathbf{J}_{G^{-1}}(\mathbf{r}) d\mathbf{r} \tag{11}$$

Evaluating the path integral in Equation 11 allows us to match every optimal choice of earnings and hours, (z^*, h^*), to a unique level of (n, α), i.e., to recover G^{-1}. As an

[2] G will be continuously differentiable as long as the utility function is twice continuously differentiable.

[3] In practice, the observed Jacobian must additionally be consistent with some function G, i.e., the Jacobian field must be conservative.

[4] We require that the set of observed (z^*, h^*) values be path connected.

example, the following parametrization of γ allows us to calculate (n, α) for any (z^*, h^*):

$$\begin{bmatrix} \log(n(z^*, h^*)) \\ \log(\alpha(z^*, h^*)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \int_{\log(z_0^*)}^{\log(z^*)} \mathbf{J}_{G^{-1}}(s, \log(h_0^*)) \begin{bmatrix} 1 \\ 0 \end{bmatrix} ds + \int_{\log(h_0^*)}^{\log(h^*)} \mathbf{J}_{G^{-1}}(\log(z^*), s) \begin{bmatrix} 0 \\ 1 \end{bmatrix} ds$$

$\square$

A.4 Proof of Proposition 2 with Kink Points

If the tax schedule is piece-wise linear, the mapping from productivities and preferences to earnings and hours worked will be more complicated for two reasons. First, bunching at kinks where the marginal tax rate increases leads to a mapping that is no longer injective. Second, if the tax schedule exhibits decreasing marginal tax rates between some tax brackets, this means that some individuals will have multiple optimal earnings levels so that $G(\log(n), \log(\alpha))$ is now a correspondence rather than a function. Bunching will mean that many types (n, α) pool on a single level of (z, h), which leads to two challenges: (1) recovering (n, α) for each bunching individual and (2) relating the levels of (n, α) for non-bunching individuals across different tax brackets. We show that (2) can be solved but (1) cannot be fixed so that we can only recover $G^{-1} : \mathcal{Z}^* \times \mathcal{H}^* \rightarrow \mathcal{N} \times \mathcal{A}$ for all individuals whose optimal earnings z^* is not a kink point of the tax schedule. (1) cannot be fixed as individuals who bunch at a kink point with the same hours of work are observationally equivalent - hence, we cannot determine (n, α) for an individual who bunches at the kink. Finally, the presence of decreasing marginal tax rates also leads to the added difficulty of relating the levels of (n, α) for non-bunching individuals across different tax brackets. However, as in the case of increasing marginal tax rates, we show that this issue can be solved.

Proof. First, we consider the case of a piece-wise linear tax schedule with two brackets with increasing marginal tax rates. Then we consider the case with two brackets and decreasing maraginal tax rates. The logic then easily extends to the case with many tax brackets, some of which may be increasing and some of which may be decreasing.

Note that all individuals (n, α) have a unique (z^*, h^*) under a piece-wise linear tax schedule with increasing rates under standard assumptions (i.e., the Hessian matrix of $U(h, e)$ is negative definite $\forall n, \alpha$). Hence, the function $G : \mathcal{N} \times \mathcal{A} \rightarrow \mathcal{Z}^* \times \mathcal{H}^*$ exists. Second, within a given tax bracket, excluding the kink points, the mapping between (n, α) and (z^*, h^*) is bijective under the assumptions in Proposition 2; this follows immediately from the proof of Proposition 2 in Appendix A.3 applied to individuals in a single tax bracket (i.e., constant tax rate). But this means that, for a given tax bracket, every (z^*, h^*) in that tax bracket corresponds to a unique (n, α). Thus, excluding the kink points of the tax schedule, every (z^*, h^*) in *every* tax bracket corresponds to a unique (n, α). Thus, the mapping between (n, α) and (z^*, h^*) is bijective globally (excluding kink points).

Now that we have established that there is a bijection between (n, α) and (z^*, h^*) $\forall z^*$ s.t. z^* is not a kink point, we need to determine how to map each (z^*, h^*) to its associated (n, α). As before, pick a particular (z_0^*, h_0^*) and normalize $(\log(n(z_0^*, h_0^*)), \log(\alpha(z_0^*, h_0^*))) = (0, 0)$. Given this normalization, we want to be able to determine the value of $(\log(n), \log(\alpha))$ that chooses any given (z^*, h^*). If z^* is in the same bracket as z_0^*, we can simply integrate the Jacobian as in the proof of Proposition 2. So consider trying to find the associated value of $(\log(n), \log(\alpha))$ for an individual with $(\log(z^*), \log(h^*))$ where z^* is in the tax bracket above z_0^* so that they are separated by a kink point at z_K.

To do this, we will first investigate the set of individuals who choose to bunch at the kink z_K and work hours h_K (there will be many different hours choices associated with z_K, we have denoted a single arbitrary choice of hours as h_K). Let the tax rate below z_K be given by T_1' and the tax rate above z_K be given by $T_2' > T_1'$. Let (n^{min}, α^{min}) denote the individual who chooses (z_K, h_K) who is just indifferent from the left (i.e., under T_1') and (n^{max}, α^{max}) denote the individual who chooses (z_K, h_K) who is just indifferent from the right (i.e., under T_2'). The individual with (n^{min}, α^{min}) satisfies the following FOCs when $z = z_K$, $h = h_K$, and $e = z_K/(n^{min} h_K)$:

$$\alpha^{min} u_c(c(z)) n^{min} e(1 - T_1') - v_h(h, e) = 0$$
$$\alpha^{min} u_c(c(z)) n^{min} h(1 - T_1') - v_e(h, e) = 0$$

The individual with (n^{max}, α^{max}) satisfies the following FOCs when $z = z_K$, $h = h_K$ and $e = z_K/(n^{max} h_K)$:

$$\alpha^{max} u_c(c(z)) n^{max} e(1 - T_2') - v_h(h, e) = 0$$
$$\alpha^{max} u_c(c(z)) n^{max} h(1 - T_2') - v_e(h, e) = 0$$

How can we relate (n^{max}, α^{max}) to (n^{min}, α^{min})? It turns out that $n^{max} = n^{min}$ and $\alpha^{max}(1 - T_2') = \alpha^{min}(1 - T_1')$ as:

$$\begin{aligned}
&\alpha^{max} u_c(c(z_K)) n^{max} \frac{z_K}{n^{max} h_K}(1 - T_2') - v_h(h_K, \frac{z_K}{n^{max} h_K}) \\
&= \alpha^{min} u_c(c(z_K)) n^{min} \frac{z_K}{n^{min} h_K}(1 - T_1') - v_h(h_K, \frac{z_K}{n^{min} h_K}) = 0
\end{aligned}$$

and

$$\alpha^{max}u_c(c(z_K))n^{max}h_K(1-T_2') - v_h(h_K, \frac{z_K}{n^{max}h_K})$$
$$= \alpha^{min}u_c(c(z_K))n^{min}h_K(1-T_1') - v_h(h_K, \frac{z_K}{n^{min}h_K}) = 0$$

Moreover, both (n^{min}, α^{min}) and (n^{max}, α^{max}) are unique.[5] Because z^* is increasing in α (by Topkis's theorem), the individuals that bunch at the kink z_K and work hours h_K are those with $n = n^{min}$ and $\alpha^{min} \leq \alpha \leq \alpha^{min}(1-T_1')/(1-T_2')$. Now, we finally show how, conditional on the normalization $(\log(n(z_0^*, h_0^*)), \log(\alpha(z_0^*, h_0^*))) = (0, 0)$, we can recover the level of (n, α) that chooses (z^*, h^*), where z^* is in the tax bracket above z_0^*. By the same logic as in the proof of Proposition 2, if γ_1 represents a curve from $(\log(z_0^*), \log(h_0^*))$ to $(\log(z_K), \log(h_K))$, we can determine the value of $(\log(n^{min}), \log(\alpha^{min}))$ by Stokes' Theorem:

$$\begin{bmatrix}\log(n^{min})\\ \log(\alpha^{min})\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} + \int_{\gamma_1} \mathbf{J}_{G^{-1}}(\mathbf{r})d\mathbf{r} \tag{12}$$

Once we know $(\log(n^{min}), \log(\alpha^{min}))$, we know $n^{max} = n^{min}$ and $\alpha^{max}(1-T_2') = \alpha^{min}(1-T_1')$. Because type (n^{max}, α^{max}) chooses $(\log(z_K), \log(h_K))$ and is just indifferent under the tax rate T_2' in the tax bracket above z_K, we can similarly apply Proposition 2 if γ_2 is a curve from $(\log(z_K), \log(h_K))$ to $(\log(z^*), \log(h^*))$:

$$\begin{bmatrix}\log(n(z^*, h^*))\\ \log(\alpha(z^*, h^*))\end{bmatrix} = \begin{bmatrix}\log(n^{max})\\ \log(\alpha^{max})\end{bmatrix} + \int_{\gamma_2} \mathbf{J}_{G^{-1}}(\mathbf{r})d\mathbf{r} = \begin{bmatrix}\log(n^{min})\\ \log(\alpha^{min}\frac{1-T_1'}{1-T_2'})\end{bmatrix} + \int_{\gamma_2} \mathbf{J}_{G^{-1}}(\mathbf{r})d\mathbf{r} \tag{13}$$

This finishes the proof in the case of two tax brackets with increasing marginal tax rates. Now we move on to the case of two tax brackets with decreasing marginal tax rates $T_2' < T_1'$. In this case, a different complication arises because individuals can have multiple optimal earnings levels - one in each tax bracket. This means that $G : \mathcal{N} \times \mathcal{A} \to \mathcal{Z}^* \times \mathcal{H}^*$ is now a correspondence instead of a function.

Note, however that $G : \mathcal{N} \times \mathcal{A} \to \mathcal{Z}^* \times \mathcal{H}^*$ is still an injective correspondence, so that no two (n, α) map to the same (z^*, h^*). Towards a contradiction, if two (n, α) were mapped to the same (z^*, h^*), then naturally both (n, α) would face the same marginal tax rate, but the assumptions in Proposition 2 ensure the relationship between (n, α) and (z^*, h^*) is bijective under a constant tax rate. Thus, G can be inverted so that $G^{-1} : \mathcal{Z}^* \times \mathcal{H}^* \to \mathcal{N} \times \mathcal{A}$ exists.

[5] Suppose not so that, for example, both $(n_1^{max}, \alpha_1^{max})$ and $(n_2^{max}, \alpha_2^{max})$ choose (z_K, h_K) and that their FOC's hold exactly under tax rate T_2'. This implies that the mapping between (n, α) and (z^*, h^*) is not bijective for individuals subject to the same tax rate, which is not possible under the assumptions in Proposition 2.

Now that we have established that $G^{-1} : \mathcal{Z}^* \times \mathcal{H}^* \to \mathcal{N} \times \mathcal{A}$ exists, we need to determine how to map each (z^*, h^*) to its associated (n, α). As before, pick a particular (z_0^*, h_0^*) and normalize $(\log(n(z_0^*, h_0^*)), \log(\alpha(z_0^*, h_0^*))) = (0, 0)$. Given this normalization, we want to be able to determine the value of $(\log(n), \log(\alpha))$ that chooses any given (z^*, h^*). If z^* is in the same bracket as z_0^*, we can simply integrate the Jacobian as in the proof of Proposition 2. So consider trying to find the associated value of $(\log(n), \log(\alpha))$ for an individual with $(\log(z^*), \log(h^*))$ where z^* is in the tax bracket above z_0^*.

The key is being able to determine a pair (z_l^*, h_l^*) and (z_h^*, h_h^*) such that some type (n_m, α_m) has two optimal earnings levels at (z_l^*, h_l^*) and (z_h^*, h_h^*). If we know that there is an individual with two optimal earnings levels at (z_l^*, h_l^*) and (z_h^*, h_h^*), we can determine (n_m, α_m) by integrating the Jacobian along a curve γ_3 between (z_0^*, h_0^*) and (z_l^*, h_l^*) as the tax rate is constant in this interval:

$$\begin{bmatrix} \log(n_m) \\ \log(\alpha_m) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \int_{\gamma_3} \mathbf{J}_{G^{-1}}(\mathbf{r}) d\mathbf{r} \tag{14}$$

Then, if we know that (z_h^*, h_h^*) is also chosen optimally by (n_m, α_m), we can simply integrate the Jacobian along a path γ_4 between (z_h^*, h_h^*) and (z^*, h^*) where z^* is in the tax bracket above z_0^* to determine the (n, α) associated with (z^*, h^*):

$$\begin{bmatrix} \log(n(z^*, h^*)) \\ \log(\alpha(z^*, h^*)) \end{bmatrix} = \begin{bmatrix} \log(n_m) \\ \log(\alpha_m) \end{bmatrix} + \int_{\gamma_4} \mathbf{J}_{G^{-1}}(\mathbf{r}) d\mathbf{r} \tag{15}$$

How can we determine such a (n_m, α_m)? For a fixed n let us consider the earnings levels chosen by each (n, α) under the tax rate in the first tax bracket T_1'. This can be determined by simply integrating $\mathbf{J}_G$ (calculated under T_1' between $(0, 0)$ and $(\log(n), \log(\alpha))$ using our normalization that $(\log(n(z_0^*, h_0^*)), \log(\alpha(z_0^*, h_0^*))) = (0, 0)$). Next, we can determine for any individual (n, α) the earnings they would choose if they faced the lower tax rate in the second bracket $T_2' < T_1'$ yet remained on the same indifference curve (denote their utility by $\bar{u}$) using the compensated elasticity (which is the elasticity of earnings with respect to one minus the tax rate, holding utility constant). This results from the following relationship, where we allow the elasticity to vary with the tax rate for a given (n, α):[6]

$$\log(z^*(T_2'; n, \alpha, \bar{u})) = \log(z^*(T_1'; n, \alpha, \bar{u})) - \int_{T_1'}^{T_2'} \xi_z^c(s; n, \alpha, \bar{u}) ds$$

More generally, we can calculate the the earnings level associated with any tax rate T'

[6]Note, we need to observe (or make assumptions about) how the earnings elasticity varies with the tax rate between T_1' and T_2' for a given individual (n, α). Naturally, assuming the compensated elasticity is constant as a function of the tax rate would be sufficient. This is a minor limitation of our approach.

on the indifference curve giving utility $\bar{u}$ as:

$$\log(z^*(T';n,\alpha,\bar{u})) = \log(z^*(T_1';n,\alpha,\bar{u})) - \int_{T_1'}^{T'} \xi_z^c(s;n,\alpha,\bar{u})ds$$

Because the compensated elasticity is always positive, we can invert this function to determine $T'(z^*;n,\alpha,\bar{u})$. One can then determine the consumption level associated with $z^*(T_2';n,\alpha,\bar{u})$ that yields utility $\bar{u}$. Define $c^*(z;n,\alpha,\bar{u})$ as the consumption level associated with a given z^* on indifference curve with utility $\bar{u}$. Noting that:

$$\left.\frac{\partial c^*}{\partial z^*}\right|_{\bar{u}} = 1 - T'(z^*;n,\alpha,\bar{u})$$

we have:

$$c^*(z^*(T_2');n,\alpha,\bar{u}) = c^*(z^*(T_1');n,\alpha,\bar{u}) + \int_{z^*(T_1')}^{z^*(T_2')} 1 - T'(s;n,\alpha,\bar{u})ds$$

Thus, for any given (n,α) who chooses $z^*(T_1')$ under the first tax rate T_1', we can determine the earnings level $z^*(T_2')$ and the corresponding consumption level $c^*(T_2')$ that would leave this individual indifferent between $z^*(T_1')$ and $z^*(T_2')$. Finally, we can determine if this $z^*(T_2')$ and the corresponding consumption level $c^*(T_2')$ lie on the actual tax schedule. If so, then this individual is indifferent between two earnings levels - one in each tax bracket. One can then search the space of (n,α) to find a (n_m,α_m) with multiple optimal earnings levels and apply Equations to 14 and 15 to map each (z^*,h^*) to the associated (n,α). This completes the case with decreasing marginal tax rates.

Note, Equations 12, 13, 14, and 15 can be easily generalized to account for more than 1 kink point, allowing us to match every (z^*,h^*) with z^* not a kink point to a unique level of (n,α) with arbitrary piece-wise linear tax schedules.

□

A.5 Proof of Proposition 3

Proof. The elasticities of $h_1,h_2,...,h_m,e_1,e_2,...,e_m$ with respect to n are related to the uncompensated elasticity and the elasticities of $h_1,h_2,...,h_m,e_1,e_2,...,e_m$ with respect to α are related to the compensated elasticities by the exact same implicit function theorem logic as in Lemma 1. More specifically, we still have:

$$\frac{\partial i^*}{\partial n}n = \frac{\partial i^*}{\partial(1-T')}(1-T')$$

and

$$\frac{\partial i^*}{\partial \alpha}\alpha = \left(\frac{\partial i^*}{\partial(1-T')} - \frac{\partial i^*}{\partial R}z^*\right)(1-T') = \frac{\partial i^*}{\partial(1-T')}\bigg|_c (1-T')$$

for $i = h_1, h_2, ..., h_m, e_1, e_2, ..., e_m$. Hence for $z = n(h_1e_1 + ... + h_me_m)$:

$$\begin{aligned} \frac{\partial z^*}{\partial n}n &= z^* + n\frac{\partial(h_1^*e_1^* + ... + h_m^*e_m^*)}{\partial n}n \\ &= z^* + n\frac{\partial(h_1^*e_1^* + ... + h_m^*e_m^*)}{\partial(1-T')}(1-T') = z^* + \frac{\partial z^*}{\partial(1-T')}(1-T') \end{aligned} \tag{16}$$

$$\begin{aligned} \frac{\partial z^*}{\partial \alpha}\alpha &= n\frac{\partial(h_1^*e_1^* + ... + h_m^*e_m^*)}{\partial \alpha}\alpha \\ &= n\frac{\partial(h_1^*e_1^* + ... + h_m^*e_m^*)}{\partial(1-T')}\bigg|_c (1-T') = \frac{\partial z^*}{\partial(1-T')}\bigg|_c (1-T') \end{aligned} \tag{17}$$

The second equality in both 16 and 17 follows by expanding the derivative according to the product rule, using the elasticity relationships term by term, and then condensing. Dividing both equations by z^* yields: $\xi_z^n = 1 + \xi_z^u$ and $\xi_z^\alpha = \xi_z^c$. Hence, our Jacobian of $G : \mathcal{N} \times \mathcal{A} \to \mathcal{Z}^* \times \mathcal{H}_1^*$ is given by:

$$\mathbf{J}_G(\log(n), \log(\alpha)) = \begin{bmatrix} \frac{\partial \log(z^*)}{\partial \log(n)} & \frac{\partial \log(z^*)}{\partial \log(\alpha)} \\ \frac{\partial \log(h_1^*)}{\partial \log(n)} & \frac{\partial \log(h_1^*)}{\partial \log(\alpha)} \end{bmatrix}(\log(n), \log(\alpha)) = \begin{bmatrix} 1 + \xi_z^u & \xi_z^c \\ \xi_{h_1}^u & \xi_{h_1}^c \end{bmatrix}(\log(n), \log(\alpha))$$

Hence, we can apply the same logic as the proof to Proposition 2 to prove the current Proposition (either in the case with a constant linear tax rate or in the more general case with a piece-wise linear tax schedule).

□

A.6 Method with Labor Supply Frictions

Labor supply frictions imply that optimal hours worked h^* and optimal earnings z^* are not equal to observed hours worked, denoted $\tilde{h}$, and observed earnings, denoted $\tilde{z}$. In order to apply Proposition 1 (or Proposition 2 in the case of heterogeneous elasticities and/or piece-wise linear tax schedules), the first step is to determine h^* and z^* from $\tilde{h}$ and $\tilde{z}$. We assume that we can recover optimal hours worked without frictions h^* (via survey, for example). Next, note that:

$$\log(z^*) = \log(\tilde{z}) + \log(h^*) - \log(\tilde{h}) + \log(e^*) - \log(\tilde{e})$$

Let us suppose that $\log(e^*) - \log(\tilde{e}) = q(\log(h^*) - \log(\tilde{h}))$ for some function $q(\cdot)$. For example, if individuals do not face any frictions in their choice of effort per hour and effort disutility is separable from hours disutility, then $\log(e^*) - \log(\tilde{e}) = 0$. Or we may assume, for example, that $\log(e^*) - \log(\tilde{e}) = \xi^c_e/\xi^c_h(\log(h^*) - \log(\tilde{h}))$. Regardless of the particular assumption, we can then infer $\log(e^*) - \log(\tilde{e})$ from $\log(h^*) - \log(\tilde{h})$, which in turn implies we can infer $\log(z^*)$. Finally, we require estimates of the elasticities of z^* and h^* with respect to the tax rate. Unfortunately, for most individuals we do not observe these elasticities (because their observed earnings and hours responses to tax rates are influenced by frictions). We can, however, observe these elasticities for subsets of the population that do not face frictions (e.g., the unemployed or Uber drivers).[7] We then assume all individuals who face frictions would have the same elasticities if they did not face any frictions as the subset of individuals who we observe without frictions.

At this point, we have all of the required elements to apply Proposition 1 (or Proposition 2 in the case of heterogeneous elasticities and/or piece-wise linear tax schedules): the distribution of optimal hours and optimal earnings as well as the elasticities of these objects with respect to the tax rate.

A.7 Dynamic Analogue to Lemma 1

Suppose that agents have made labor supply decisions up to some time t, so that their human capital K and past labor supply decisions at times $1, ..., t-1$ are fixed. We want to show that the relationships $\xi^{n_t}_{z_t} = 1 + \xi^u_{z_t}$, $\xi^{n_t}_{h_t} = \xi^u_{h_t}$, $\xi^\alpha_{z_t} = \xi^c_{z_t}$, and $\xi^\alpha_{h_t} = \xi^c_{h_t}$ hold. Let us denote the growth rate of the effort wage at time t as $q_t(h_t, e_t)$ and the cumulative growth $Q_t \equiv \prod_{s=1}^{t-1} q_t(h_t, e_t)$. The problem for the individual starting at a time t can be represented as (using the fact that for any time $s \geq t$, $n_s = n_0 K Q_s = n_t \prod_{k=t}^{s-1} q_k(h_k, e_k) = n_t(Q_s/Q_t)$):

$$\max_{\{h\}_{s=t}^L, \{e\}_{s=t}^L} \sum_{s=t}^{L} \beta^s \left[\alpha u(c_s) - v(h_s, e_s)\right]$$

$$\text{s.t. } c_s \leq n_t \frac{Q_s}{Q_t} h_s e_s (1 - T') + R$$

Alternatively, we could define $\nu = n_t(1 - T')$ rewrite this problem as:

$$\max_{\{h\}_{s=t}^L, \{e\}_{s=t}^L} \sum_{s=t}^{L} \beta^s \left[\alpha u(c_s) - v(h_s, e_s)\right]$$

$$\text{s.t. } c_s \leq \nu \frac{Q_s}{Q_t} h_s e_s + R$$

[7] Alternatively, one could try to estimate these elasticities by surveying individuals' choices of earnings and hours worked under different tax rates.

Note then that for any choice variable $i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}$, we have that:

$$\frac{\partial \log(i^*)}{\partial \log(n_t)} = \frac{\partial \log(i^*)}{\partial \log(\nu)} \frac{\partial \log(\nu)}{\partial \log(n_t)} = \frac{\partial \log(i^*)}{\partial \log(\nu)} = \frac{\partial \log(i^*)}{\partial \log(\nu)} \frac{\partial \log(\nu)}{\partial \log(1-T')} = \frac{\partial \log(i^*)}{\partial \log(1-T')}$$

Hence, setting $i = h_t$ immediately gives us $\xi_{h_t}^{n_t} = \xi_{h_t}^{u}$. Similarly, we have $\xi_{e_t}^{n_t} = \xi_{e_t}^{u}$; since $\log(z_t^*) = \log(n_t) + \log(h_t^*) + \log(e_t^*)$, we get that $\xi_{z_t}^{n_t} = 1 + \xi_{z_t}^{u}$. Next, suppose that we take first order conditions with respect to choice variables h_k and e_k (hours and effort per hour at arbitrary time k), recalling $z_s = n_t(Q_s/Q_t)h_s e_s$:

$$\sum_{s=t}^{L} \beta^s \left[\alpha u'(c_s^*) \frac{\partial z_s}{\partial h_k}(1-T')\right] - \beta^k v_1(h_k^*, e_k^*) = 0$$

$$\sum_{s=t}^{L} \beta^s \left[\alpha u'(c_s^*) \frac{\partial z_s}{\partial e_k}(1-T')\right] - \beta^k v_2(h_k^*, e_k^*) = 0$$

Note that in the above FOCs, $\partial z_s/\partial h_k$ and $\partial z_s/\partial e_k$ are functions that are evaluated at the optimal choices $\{h^*\}_{s=t}^{L}, \{e^*\}_{s=t}^{L}$ (but we omit these arguments for the sake of brevity). Defining $\theta = \alpha(1-T')$, for $i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}$ we can rewrite our FOCs as:

$$\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial h_k}\right] - \beta^k v_1(h_k, e_k) = 0$$

$$\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial e_k}\right] - \beta^k v_2(h_k^*, e_k^*) = 0$$

These first order conditions allow us to derive $\partial \log(i^*)/\partial \log(\alpha)$ using the implicit function theorem. Using the fact that α does not enter the FOCs except through its affect on θ, we have:

$$\frac{\partial \log(i^*)}{\partial \log(\alpha)} = \frac{\partial \log(i^*)}{\partial \log(\theta)} \frac{\partial \log(\theta)}{\partial \log(\alpha)} = \frac{\partial \log(i^*)}{\partial \log(\theta)}$$

Moreover, we also have that:

$$\frac{\partial \log(i^*)}{\partial \log(1-T')} = \frac{\partial \log(i^*)}{\partial \log(\theta)} \frac{\partial \log(\theta)}{\partial \log(1-T')} + \left.\frac{\partial \log(i^*)}{\partial \log(1-T')}\right|_{\theta} = \frac{\partial \log(i^*)}{\partial \log(\theta)} + \left.\frac{\partial \log(i^*)}{\partial \log(1-T')}\right|_{\theta}$$

Thus:

$$\frac{\partial \log(i^*)}{\partial \log(\alpha)} = \frac{\partial \log(i^*)}{\partial \log(1-T')} - \left.\frac{\partial \log(i^*)}{\partial \log(1-T')}\right|_{\theta}$$

We now are going to show that:

$$\left.\frac{\partial \log(i^*)}{\partial \log(1-T')}\right|_{\theta} = \sum_{j=t}^{L} \frac{\partial \log(i^*)}{\partial R_j} z_j^*(1-T')$$

Differentiating the FOCs with respect to $1-T'$, holding θ constant, the implicit function theorem gives us the following two relationships (note there will be two such equations for each time k):

$$\begin{aligned}
&\sum_{s=t}^{L} \beta^s \left[\theta u''(c_s^*) \frac{\partial z_s}{\partial h_k} z_s^*(1-T')\right] \\
&+ \sum_{i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}} \left.\frac{\partial i^*}{\partial \log(1-T')}\right|_{\theta} \frac{\partial}{\partial i} \left(\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial h_k}\right] - \beta^k v_1(h_k^*, e_k^*)\right) = 0
\end{aligned} \tag{18}$$

$$\begin{aligned}
&\sum_{s=t}^{L} \beta^s \left[\theta u''(c_s^*) \frac{\partial z_s}{\partial e_k} z_s^*(1-T')\right] \\
&+ \sum_{i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}} \left.\frac{\partial i^*}{\partial \log(1-T')}\right|_{\theta} \frac{\partial}{\partial i} \left(\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial e_k}\right] - \beta^k v_2(h_k^*, e_k^*)\right) = 0
\end{aligned} \tag{19}$$

Next, we define $\partial i^*/\partial R_j$ as the derivative of i^* with respect to an income shock in period j. Differentiating the FOCs with respect to R_j and multiplying by $z_j^*(1-T')$ gives us:

$$\begin{aligned}
&\beta^j \left[\theta u''(c_j^*) \frac{\partial z_j}{\partial h_k} z_j^*(1-T')\right] \\
&+ \sum_{i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}} \frac{\partial i^*}{\partial R_j} z_j^*(1-T') \frac{\partial}{\partial i} \left(\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial h_k}\right] - \beta^k v_1(h_k^*, e_k^*)\right) = 0
\end{aligned}$$

$$\begin{aligned}
&\beta^j \left[\theta u''(c_j^*) \frac{\partial z_j}{\partial e_k} z_j^*(1-T')\right] \\
&+ \sum_{i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}} \frac{\partial i^*}{\partial R_j} z_j^*(1-T') \frac{\partial}{\partial i} \left(\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial e_k}\right] - \beta^k v_2(h_k^*, e_k^*)\right) = 0
\end{aligned}$$

Summing these FOCs over j from t to L and switching the index of summation from j to s in the first term, we get:

$$\begin{aligned}
&\sum_{s=t}^{L} \beta^s \left[\theta u''(c_s^*) \frac{\partial z_s}{\partial h_k} z_s^*(1-T')\right] \\
&+ \sum_{i \in \{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}} \sum_{j=t}^{L} \frac{\partial i^*}{\partial R_j} z_j^*(1-T') \frac{\partial}{\partial i} \left(\sum_{s=t}^{L} \beta^s \left[\theta u'(c_s^*) \frac{\partial z_s}{\partial h_k}\right] - \beta^k v_1(h_k^*, e_k^*)\right) = 0
\end{aligned} \tag{20}$$

$$
\begin{aligned}
&\sum_{s=t}^{L}\beta^s\left[\theta u''(c_s^*)\frac{\partial z_s}{\partial e_k}z_s^*(1-T')\right] \\
(21)\quad &+\sum_{i\in\{h\}_{s=t}^{L},\{e\}_{s=t}^{L}}\sum_{j=t}^{L}\frac{\partial i^*}{\partial R_j}z_j^*(1-T')\frac{\partial}{\partial i}\left(\sum_{s=t}^{L}\beta^s\left[\theta u'(c_s^*)\frac{\partial z_s}{\partial e_k}\right]-\beta^k v_2(h_k^*,e_k^*)\right)=0
\end{aligned}
$$

Matching terms in Equations 20 and 21 with Equations 18 and 19 as in Lemma 1 (recognizing that these equations hold for all time periods k) we can state that:

$$\left.\frac{\partial i^*}{\partial\log(1-T')}\right|_\theta=\sum_{j=t}^{L}\frac{\partial i^*}{\partial R_j}z_j^*(1-T')$$

Dividing by i^* yields:

$$\left.\frac{\partial\log(i^*)}{\partial\log(1-T')}\right|_\theta=\sum_{j=t}^{L}\frac{\partial\log(i^*)}{\partial R_j}z_j^*(1-T')$$

Thus, we have that:

$$\frac{\partial\log(i^*)}{\partial\log(\alpha)}=\frac{\partial\log(i^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(i^*)}{\partial R_j}z_j^*(1-T')$$

Hence:

$$\frac{\partial\log(h_t^*)}{\partial\log(\alpha)}=\frac{\partial\log(h_t^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(h_t^*)}{\partial R_j}z_j^*(1-T')$$

and

$$\frac{\partial\log(e_t^*)}{\partial\log(\alpha)}=\frac{\partial\log(e_t^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(e_t^*)}{\partial R_j}z_j^*(1-T')$$

Using $\log(z_t^*)=\log(n_t)+\log(h_t^*)+\log(e_t^*)$ we get:

$$\frac{\partial\log(z_t^*)}{\partial\log(\alpha)}=\frac{\partial\log(z_t^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(z_t^*)}{\partial R_j}z_j^*(1-T')$$

Defining the compensated elasticities in the dynamic setting to be equal to:

$$\xi_{h_t}^c\equiv\frac{\partial\log(h_t^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(h_t^*)}{\partial R_j}z_j^*(1-T')$$

and

$$\xi_{z_t}^c\equiv\frac{\partial\log(z_t^*)}{\partial\log(1-T')}-\sum_{j=t}^{L}\frac{\partial\log(z_t^*)}{\partial R_j}z_j^*(1-T')$$

we have our stated relationship that $\xi^{\alpha}_{h_t} = \xi^{c}_{h_t}$ and $\xi^{\alpha}_{z_t} = \xi^{c}_{z_t}$ as desired. In the dynamic case, the compensated elasticity represents how individuals respond to a change in marginal tax rates less the lifetime income effects that occur due to this change in the tax rate today as well as in all future periods.

The key idea is still that changing the tax rate leads to both a substitution effect as well as an income effect. The difference in the dynamic setting is that the income effect of a tax change yields an income boost not only in the current period but also in future periods (because tax changes are permanent). Because α still only causes a substitution effect, to relate changes in α to changes in the tax rate, we need to net out both current and future income effects, leading to a modified compensated elasticity in the dynamic setup. Note that perfectly estimating the lifetime income effects of tax changes may be empirically challenging as it requires us to both estimate future earnings z_j^* as well as current responses to current and future income shocks $\partial \log(i^*)/\partial R_j$ for $j = t, t+1, ..., L$. Nonetheless, we expect that we can make some sensible assumptions on these terms so as to apply our method even when productivities are determined by previous labor supply decisions.

A.8 Dynamic Case with Savings

We augment the discussion from Section IV.C to include savings. Suppose that individuals can save at interest rate $1 + r$ and choose a level of assets a_t each period:

$$\begin{aligned}
&\max_{\{h\}_{t=0}^{L}, \{e\}_{t=0}^{L}, \{a\}_{t=0}^{L}, K} \sum_{t=0}^{L} \beta^t \left[\alpha u(c_t) - v(h_t, e_t)\right] - \kappa(K) \\
&\text{s.t. } c_t \leq n_0 K Q_t h_t e_t (1 - T') + R + (1+r) a_{t-1} - a_t \\
&a_L = 0
\end{aligned}$$

Suppose that agents have made labor supply decisions up to some time t, so that their human capital K and past labor supply decisions at times $1, ..., t-1$ are fixed. The problem for the individual starting at a time t can be represented as (using the fact that for any time $s \geq t$, $n_s = n_0 K Q_s = n_0 K Q_t \prod_{k=t}^{s-1} q_k(h_k, e_k) = n_t \prod_{k=t}^{s-1} q_k(h_k, e_k)$):

$$\begin{aligned}
&\max_{\{h\}_{s=t}^{L}, \{e\}_{s=t}^{L}, \{a\}_{s=t}^{L}} \sum_{s=t}^{L} \beta^s \left[\alpha u(c_s) - v(h_s, e_s, K)\right] \\
&\text{s.t. } c_s \leq n_s h_s e_s (1 - T') + R + (1+r) a_{s-1} - a_s \\
&a_L = 0
\end{aligned}$$

From the perspective of a single time period t, there are three relevant pieces of heterogeneity: the MRS α, the effort wage $n_t = n_0 K Q_t$, and the level of available savings $\sigma_t = (1+r)a_{t-1}$. If we can observe earnings, hours worked, and savings we can recover the function G that maps each $(\log(n_t), \log(\alpha), \sigma_t)$ to $(\log(z_t^*), \log(h_t^*), \sigma_t)$. Denote $\theta_i^{\sigma_t} \equiv \partial \log(i^*)/\partial\sigma_t = \partial \log(i^*)/\partial R_t$, the one-time income effect semi-elasticity (which could be empirically estimated as the behavioral response to a one-time income shock). Using the dynamic version of Lemma 1 discussed in Appendix A.7 (which still holds with savings, as the additional first order conditions for a_s do not change the relationship between elasticities with respect to n and α and the tax rate)[8] the Jacobian of this function is given by:

$$\begin{aligned}
\mathbf{J}_G(\log(n_t), \log(\alpha), \sigma_t) &= \begin{bmatrix} \frac{\partial \log(z_t^*)}{\partial \log(n_t)} & \frac{\partial \log(z_t^*)}{\partial \log(\alpha)} & \frac{\partial \log(z_t^*)}{\partial \sigma_t} \\ \frac{\partial \log(h_t^*)}{\partial \log(n_t)} & \frac{\partial \log(h_t^*)}{\partial \log(\alpha)} & \frac{\partial \log(h_t^*)}{\partial \sigma_t} \\ \frac{\partial \sigma_t}{\partial \log(n_t)} & \frac{\partial \sigma_t}{\partial \log(\alpha)} & \frac{\partial \sigma_t}{\partial \sigma_t} \end{bmatrix} \\
&= \begin{bmatrix} 1+\xi_{z_t}^u & \xi_{z_t}^c & \theta_{z_t}^{\sigma_t} \\ \xi_{h_t}^u & \xi_{h_t}^c & \theta_{h_t}^{\sigma_t} \\ 0 & 0 & 1 \end{bmatrix} (\log(n_t), \log(\alpha), \sigma_t)
\end{aligned}$$

The mapping G is bijective under the same conditions as in Proposition 2 as all principal minors of $\mathbf{J}_G(\log(n_t), \log(\alpha), \sigma_t)$ are positive. So if we can observe z_t^*, h_t^*, and σ_t, along with the elasticities to form $\mathbf{J}_G$, we can recover G^{-1} by the same process as in the proof of Proposition 2. Note that if $u(c)$ is linear in consumption so that income effects are 0, we can identify G^{-1} without observing σ_t as σ_t will not affect optimal choice of earnings or hours worked.[9]

A.9 Heterogeneity in Unearned Income

Suppose individuals have heterogeneity in unearned income M, so that the individual problem is:

$$\begin{aligned}
&\max_{h,e} \ \alpha u(c) - v(h,e) \\
&\text{s.t. } c \leq nhe(1-T') + R + M
\end{aligned}$$

Suppose further that we can observe unearned income $M \in \mathcal{M}$ and that we want to recover the function that maps $(\log(z^*), \log(h^*), M)$ to $(\log(n), \log(\alpha), M)$, denoted $G^{-1}: \mathcal{Z}^* \times \mathcal{H}^* \times \mathcal{M} \to \mathcal{N} \times \mathcal{A} \times \mathcal{M}$. Defining $\phi_i = \partial \log(i^*)/\partial M$, the Jacobian matrix is now

[8] This proof is omitted as it is contains no new insights beyond the dynamic analogue in Section A.7.

[9] This can be seen by inverting $\mathbf{J}_G$ noting that $\theta_{z_t}^{\sigma_t} = \theta_{h_t}^{\sigma_t} = 0$.

given by:

$$\mathbf{J}_G(\log(n), \log(\alpha), M) = \begin{bmatrix} \frac{\partial \log(z^*)}{\partial \log(n)} & \frac{\partial \log(z^*)}{\partial \log(\alpha)} & \frac{\partial \log(z^*)}{\partial M} \\ \frac{\partial \log(h^*)}{\partial \log(n)} & \frac{\partial \log(h^*)}{\partial \log(\alpha)} & \frac{\partial \log(h^*)}{\partial M} \\ \frac{\partial M}{\partial \log(n)} & \frac{\partial M}{\partial \log(\alpha)} & \frac{\partial M}{\partial M} \end{bmatrix}$$

$$= \begin{bmatrix} 1+\xi_z^u & \xi_z^c & \phi_z \\ \xi_h^u & \xi_h^c & \phi_h \\ 0 & 0 & 1 \end{bmatrix} (\log(n), \log(\alpha), M)$$

This matrix has positive principal minors under the same conditions as in Proposition 2 (hence G is globally invertible); the rest of the procedure to recover G^{-1} is unchanged from the proof of Proposition 2. Essentially, if individuals differ in terms of unearned income, we first need to subtract out the component of optimal hours and optimal earnings due to income effects using the parameters ϕ_h and ϕ_z. Then, we can recover n and α from the component of optimal earnings and optimal hours that is not due to unearned income effects.

A.10 Non-Separable Utility

It is useful to consider how our assumption of separable utility affects our result. Suppose we have a utility function as follows:

$$\max_{h,e} \; u(\alpha c, h, e)$$
$$\text{s.t. } c \leq nhe(1-T') + R$$

Using the exact same sort of arguments as in Appendix 1 to prove Lemma A.1, we can show that the Jacobian matrix of $G : \mathcal{N} \times \mathcal{A} \to \mathcal{Z}^* \times \mathcal{H}^*$ is now:

$$\mathbf{J}_G(\log(n), \log(\alpha)) = \begin{bmatrix} \frac{\partial \log(z^*)}{\partial \log(n)} & \frac{\partial \log(z^*)}{\partial \log(\alpha)} \\ \frac{\partial \log(h^*)}{\partial \log(n)} & \frac{\partial \log(h^*)}{\partial \log(\alpha)} \end{bmatrix} = \begin{bmatrix} 1+\xi_z^u & \xi_z^c + \frac{\partial \log(z^*)}{\partial R} c(z^*) \\ \xi_h^u & \xi_h^c + \frac{\partial \log(h^*)}{\partial R} c(z^*) \end{bmatrix} (\log(n), \log(\alpha))$$

We can still recover G^{-1} using the method of Proposition 2 as long as this new Jacobian matrix has positive principal minors. Sufficient conditions for this are: $1+\xi_z^u > 0$, $\xi_h^c + (\partial \log(h^*)/\partial R) c(z^*) > 0$ and $(1+\xi_z^u)\,(\xi_h^c + (\partial \log(h^*)/\partial R) c(z^*)) > (\xi_z^c + (\partial \log(z^*)/\partial R) c(z^*))\,\xi_h^u$. These conditions will hold as long as income effects are not too large.

B Data Appendix

B.1 ATUS Data Description

The American Time Use Survey (ATUS) is an annual repeated cross-sectional survey conducted on a subset of individuals who have participated in the CPS. We have data for individuals surveyed in the years 2003-2015 (individuals are only surveyed once). In addition to earnings data, the ATUS asks respondents to meticulously detail all of their activities on a particular (random) "diary day".

B.1.1 Sample Construction

We assume that the noisy "diary day" measure of hours worked is representative of this individual's average daily hours worked. We implicitly assume that all individuals work Monday-Friday, thereby dropping individuals whose randomly assigned diary day happened to fall on a Saturday or Sunday. Moreover, because we only have information on individuals' earnings in their primary occupation, we drop all individuals who have ≥ 2 jobs; this is around 3.4% of people. We also do not observe days worked per year, so we impute that all individuals work 250 days a year unless they report being part-time individuals and work > 8 hours on their diary day, in which case we impute their days worked as 125. In other words, we assume that part-time individuals who work long hours (> 8 hours per day) only work half of the usual working days. However, this only applies to a small number of individuals as full-time workers comprise 85% of our sample. We keep all individuals that have positive earnings in our sample, abstracting from the possibility of joint familial labor supply decisions - our findings all hold with the smaller sample of single individuals, shown in Appendix C.3. We drop individuals who say they are involuntarily part-time in the CPS Annual Social and Economic Supplement (ASEC); this hopefully mitigates the effect of labor supply frictions. However, because we can only match around 30% of our ATUS sample to the CPS ASEC, there are 4,413 part-time individuals for whom we do not know whether they are involuntarily part-time.[10] Because nearly 80% of the matched part-time individuals do not report to be involuntarily part-time, we keep the non-matched part-time individuals in our sample. However, our findings are robust to dropping all part-time workers that we cannot match to the CPS and all part-time workers that we can match who report to be involuntarily part-time. Lastly, we winsorize the top and bottom 0.5% of hourly wage earners. Our final sample from the ATUS then consists of data on (inflation adjusted) earnings and diary hours worked for 35,004 unique individuals from the years 2003-2015.

[10]While the ATUS is a sub-sample of the CPS, the linking variables in the CPS ASEC only allow us to uniquely identify a subset of households in the ATUS.

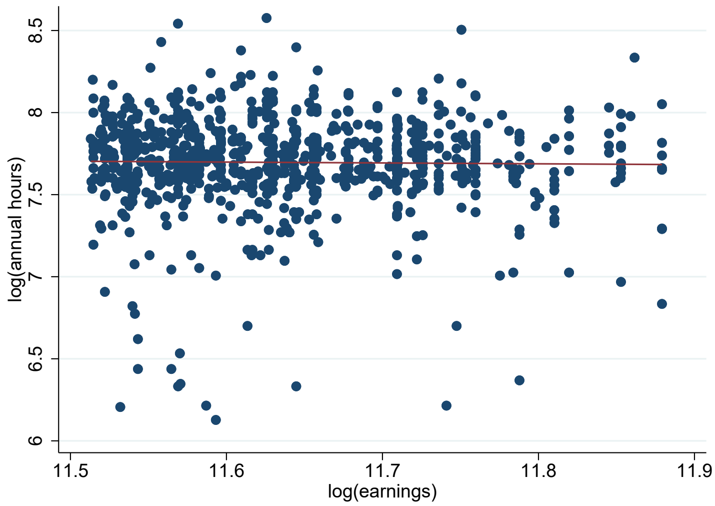


Figure 1: $\log(AnnualHours)$ vs. $\log(Earnings)$, Earnings Above \$100,000 in ATUS

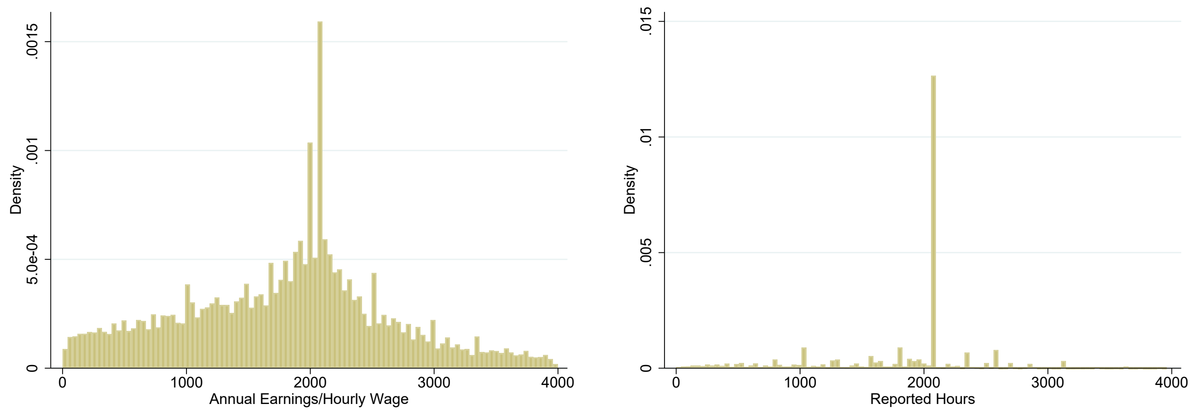
B.1.2 Top Coding Earnings

The ATUS top-codes individual wage earnings at $\approx \$145,000$ (the top-coding is not inflation adjusted so does not change over time). To deal with this, we assume that annual hours (which we do observe for top-coded individuals) and earnings are independent at the highest earnings levels. This allows us to simulate the earnings of these individuals by drawing from a Pareto distribution (with Pareto parameter 2), which matches the observed top earnings distribution quite well, see Saez (2001). In support of this independence assumption, Figure 1 illustrates a near zero correlation between earnings and annual hours worked for individuals who are not top-coded and make above $\$100,000$ per year.

B.2 CPS Hours Worked Measure

The CPS Annual Social and Economic Supplement (ASEC), which has data on individual earnings, also asks people how many hours they typically work per week as well as the number of weeks they work per year. This may seem like a natural data source for our purposes; however, we believe the measure of hours worked in this dataset suffers from substantial reporting error. Individuals appear to report "notional" hours of work, which may be drastically different from the number of hours they actually work. To support this assertion, we examine how reported hours of work in the CPS compares to actual hours worked for hourly wage workers, a subset of individuals for whom we believe we can reasonably accurately measure their actual hours worked by dividing annual earnings by their hourly wage rate.[11] Figure 2 plots annual hours worked for hourly wage workers

[11]This measure is still imperfect due to overtime and bonuses.



(a) Annual Earnings / Hourly Wage

(b) Reported Hours

Figure 2: Hours Worked in the CPS

only: Panel 2a plots annual hours worked, calculated as wage earnings divided by hourly wage and Panel 2b plots reported annual hours (reported hours per week multiplied by reported weeks per year). In particular, 48% of hourly wage workers report working 40 hours per week and 52 weeks per year.[12] This is clearly not in alignment with their observed hours worked, calculated using their earnings divided by the wage rate; hence we conclude that the hours worked measure from the CPS is a poor indicator of actual hours worked for hourly workers. Because the reported annual hours worked distribution is similar for non-hourly workers, we strongly suspect the same reporting bias plagues the distribution of annual hours worked for non-hourly workers in the CPS.

Conversely, the measure of hours worked from the ATUS seems to match relatively well with the distribution of actual hours worked for the hourly wage workers in the CPS. We use the 1,253 hourly workers in the ATUS who can be matched in the CPS.[13] For this set of workers, Figure 3 compares the (kernel smoothed) distributions of annual hours worked constructed using the (a) diary day method and (b) annual earnings divided by hourly wage from the CPS (as shown above in Figure 2a). Despite a sample of only around a thousand individuals, these distributions are relatively similar, providing suggestive evidence that the ATUS diary day measure is giving us a noisy, yet relatively unbiased, estimate of hours worked. The ATUS density has slightly more pronounced peaks at ≈ 1000 hours and ≈ 2000 hours simply due to the fact that we multiply diary day hours by 250 for full-time workers and 125 for part-time workers who work > 8 hours per day.

[12]Individuals are likely reporting weeks employed as opposed to working weeks, which would net out vacation.

[13]While the ATUS is a sub-sample of the CPS, the linking variables in the CPS ASEC only allow us to uniquely identify a subset of households in the ATUS.

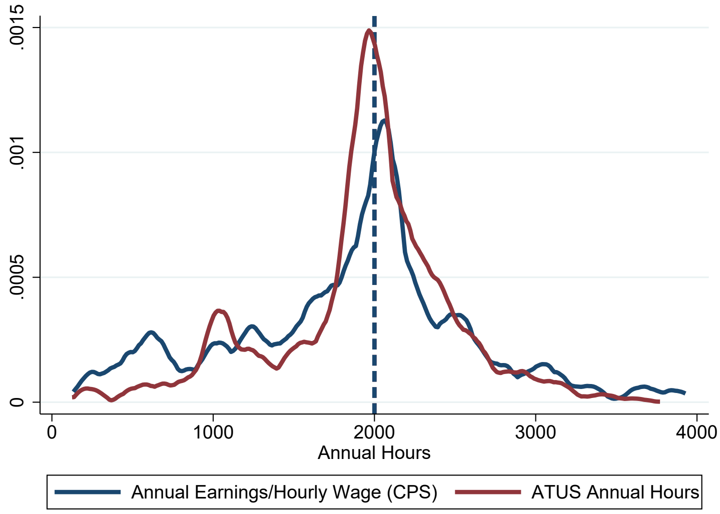


Figure 3: Annual Hours: CPS vs. ATUS (Hourly Workers)

C Additional Analysis and Results Appendix

C.1 Elasticity Heterogeneity

We augment the analysis from Section V to allow for heterogeneity in elasticities across the space of hours worked. The median elasticity is still assumed to be $\xi_z^c = \xi_h^c = 0.15$ and income effects are 0. We explore two scenarios: (1) elasticities linearly increase in log hours so that the lowest hours-worked individual in society has an elasticity around 0 and the highest hours worked individual has an elasticity around 0.2; and (2) elasticities linearly decrease in log hours so that the lowest hours worked individual has an elasticity around 0.6 and the highest hours-worked individual has an elasticity around 0. Allowing elasticities to vary with hours requires us to integrate the inverse Jacobian matrix as in Proposition 2. Given that $\xi_h^c = a + b\log(h)$ for some constants a, b we get:

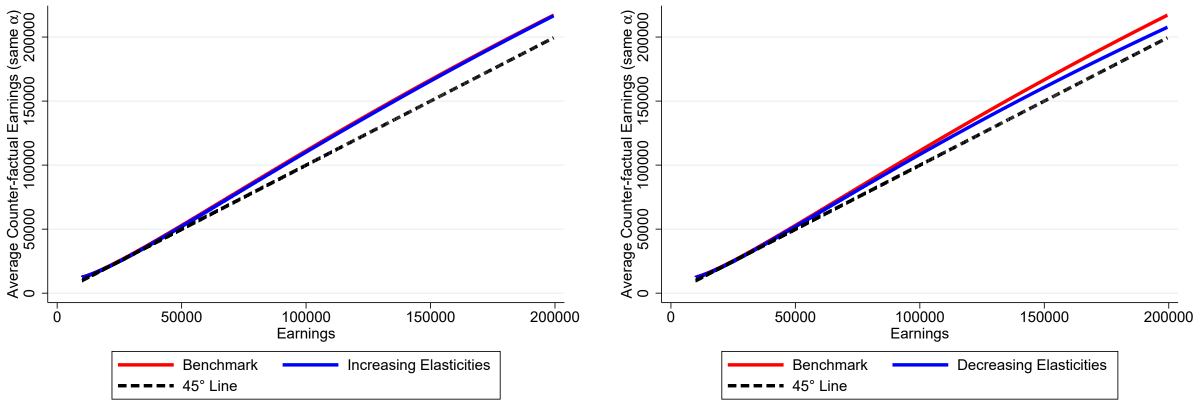
$$\mathbf{J}_{G^{-1}} == \begin{bmatrix} 1 & -1 \\ -1 & 1 + \frac{1}{a+b\log(h)} \end{bmatrix}$$

Hence, to calculate $(\log(n), \log(\alpha))$ that optimally chooses $(\log(z^*), \log(h^*))$ we integrate this inverse Jacobian along a path between $(\log(z_0), \log(h_0))$ and $(\log(z^*), \log(h^*))$:

$$\log(n) = \int_{\log(z_0)}^{\log(z^*)} 1 ds + \int_{\log(h_0)}^{\log(h^*)} -1 ds$$

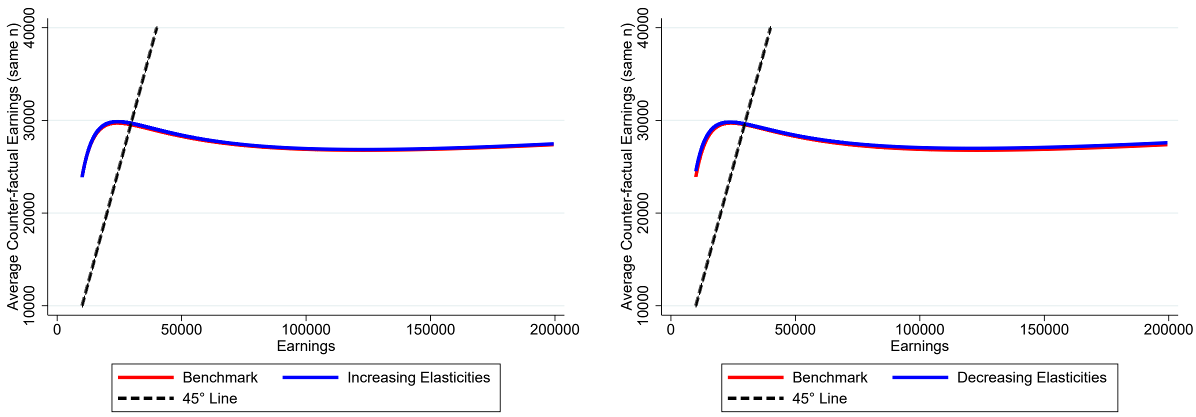
$$\log(n) = \int_{\log(z_0)}^{\log(z^*)} -1 ds + \int_{\log(h_0)}^{\log(h^*)} 1 + \frac{1}{a + b\log(h)} ds$$

Once we know $(\log(n), \log(\alpha))$ that optimally chooses each $(\log(z^*), \log(h^*))$, we can invert this function to determine $(\log(z^*), \log(h^*))$ as a function of $(\log(n), \log(\alpha))$. This function then allows us to determine counter-factual incomes. We plot average counter-factual earnings (if all individuals had the same α) by actual earnings levels in Figure 4 for both increasing elasticities and decreasing elasticities. We plot average counter-factual earnings (if all individuals had the same n) by actual earnings levels in Figure 5 for both increasing elasticities and decreasing elasticities. As discussed in the main text, these Figures look nearly identical to the benchmark case without elasticity heterogeneity. This results because, while the elasticities vary with hours worked, the correlation between hours worked and earnings is not overly strong so that average elasticities do not vary substantially over the earnings distribution.



(a) Increasing Elasticity in Hours: $\frac{d\xi_h^c}{dh} > 0$ (b) Decreasing Elasticity in Hours: $\frac{d\xi_h^c}{dh} < 0$

Figure 4: Average Counter-Factual Earnings (same α), Heterogeneous Elasticities



(a) Increasing Elasticity in Hours: $\frac{d\xi_h^c}{dh} > 0$ (b) Decreasing Elasticity in Hours: $\frac{d\xi_h^c}{dh} < 0$

Figure 5: Average Counter-Factual Earnings (same n), Heterogeneous Elasticities

C.2 Labor Supply Frictions

The National Study of the Changing Workforce (NSCW) has data not only on earnings and hours worked, but also on the number of hours each individual would prefer to work if they faced no frictions for 1992, 1997, 2002, and 2008; we pool all of these years to construct our sample and adjust earnings for inflation. The NSCW asks individuals "If you could do what you wanted to do, ideally how many hours in total would you like to work each week?" We use the response to this question as our measure of optimal hours of work. The distribution of actual and ideal weekly hours worked from the NSCW is shown in Figure 6. The distributions of actual and ideal weekly hours are somewhat similar although, on average, ideal hours is about 10% less than actual hours.

From Section IV.B, we also need elasticities of optimal earnings and optimal hours with respect to the tax rate. We use estimates of earnings and hours elasticities for self-employed people, who are likely subject to far fewer frictions than the non-self-employed.[14] We use estimates from Heim (2010) who finds that the real (as opposed to reported) earnings elasticity with respect to the tax rate for the self-employed is 0.4, i.e., $\xi_z^u = 0.4$. We assume that effort is exogenous so that $\xi_h^u = 0.4$.[15] Finally, we assume income effects are 0 so that $\xi_z^u = \xi_z^c = \xi_h^u = \xi_h^c = 0.4$.

First, we calculate the distribution of optimal hours worked and optimal earnings using data on optimal hours worked. Optimal earnings are equal to $z^* = nh^*$ where productivity $n = \tilde{z}/\tilde{h}$ is recovered by dividing observed earnings $\tilde{z}$ by observed hours worked $\tilde{h}$. We then use the distribution of (z^*, h^*) to get the distribution of productivities and preferences exactly as in Proposition 1, using our elasticity estimates for individuals who face no frictions to form the Jacobian used to construct the inverse function. Next, we determine the counter-factual optimal hours worked for each individual if they had wage n_0: $h^*(n_0, \alpha)$. We then use the distribution of hours frictions for individuals who, in actuality, have productivity $\approx n_0$ and preferences $\approx \alpha$.[16] For all individuals, we draw a friction $\epsilon_{n_0,\alpha}$ from this distribution of frictions for individuals whose productivity is $\approx n_0$ and preferences are $\approx \alpha$ and add it to their counterfactual hours. Then our counter-factual earnings level for each person is given by $z_{n_0}^{CF,Frictions} = n_0(h^*(n_0, \alpha) + \epsilon_{n_0,\alpha})$. Note that because the value of $\epsilon_{n_0,\alpha}$ is random, the value of $z_{n_0}^{CF,Frictions}$ is different even for individuals with the same (n, α). An analogous process is used to calculate the counter-factual distribution assuming all individuals have preferences α_0.

[14] Note, we assume that individuals who are not self-employed would respond to taxes in the same manner as self-employed people if they did not face labor market frictions.

[15] We can, however, implement our method with frictions if effort is endogenous if we make an assumption on the relationship between effort frictions and hours frictions, see Section IV.B.

[16] More precisely, we split the distribution of n and α into deciles and sample from the partition containing n_0 and α.

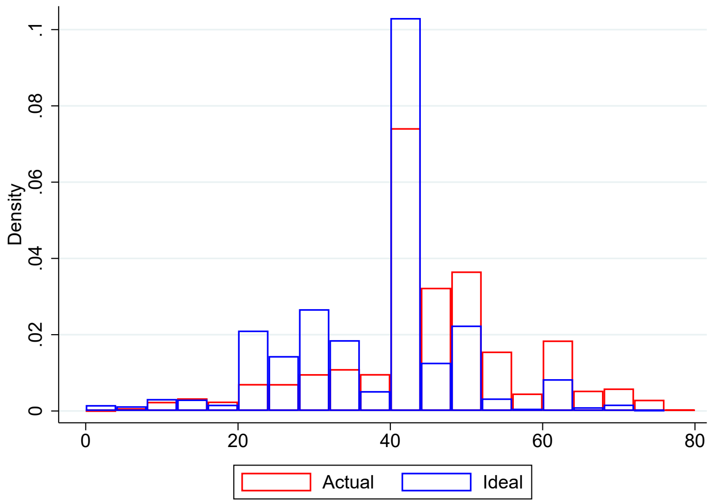
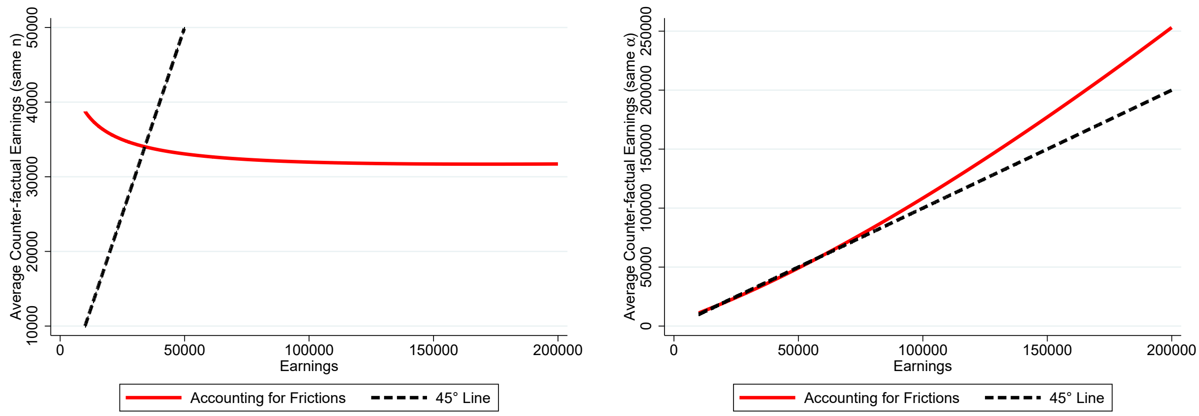


Figure 6: Actual vs. Ideal Hours Worked per Week, NSCW



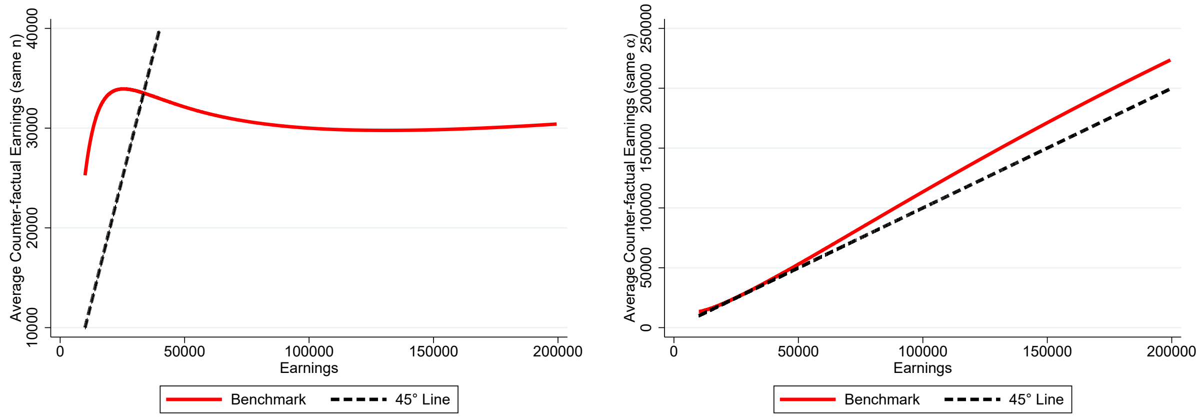
(a) Homogeneous n

(b) Homogeneous α

Figure 7: Average Counter-Factual Earnings, Accounting for Labor Market Frictions

Note: Elasticities used to construct this figure are from Heim (2010): $\xi_z^u = \xi_z^c = \xi_h^u = \xi_h^c = 0.4$.

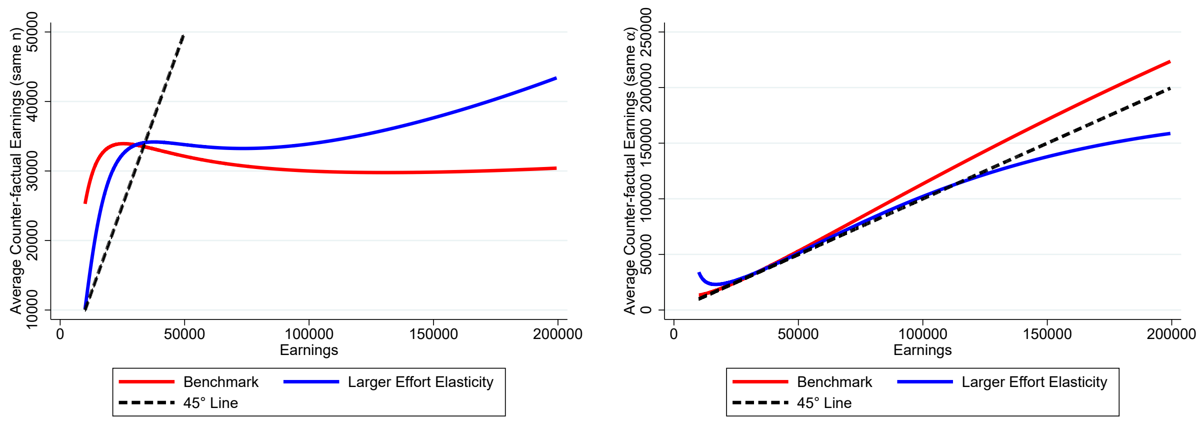
In Figure 7 we show the average counter-factual earnings level vs. actual earnings assuming all individuals had the same n (7a) and assuming all individuals had the same α (7b). First, note that accounting for frictions has a relatively modest effect on the average counter-factual earnings plots relative to the Figures shown in Section V even though we use the NSCW instead of the ATUS and we use elasticities from Heim (2010) instead of Chetty (2012). Figure 7 is qualitatively similar to Figure 3 - the average counter-factual earnings curve in Figure 7a is relatively flat (compared to the 45° line) and high earnings individuals have lower average counter-factual earnings than middle earnings individuals, implying that higher earnings people have weaker preferences for consumption. In Figure 7b, high earnings individuals have higher average counter-factual earnings than in actuality, again suggesting that they have weaker preferences for consumption.



(a) Homogeneous n (b) Homogeneous α

Figure 8: Average Counter-Factual Earnings for Singles, Benchmark Elasticities

Note: Elasticities used to construct this figure are from Chetty (2012): $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$.



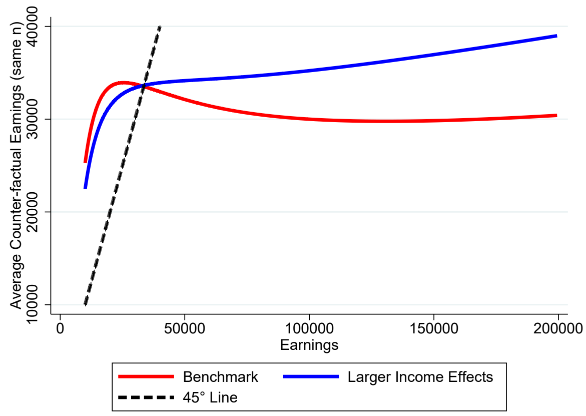
(a) Homogeneous n (b) Homogeneous α

Figure 9: Average Counter-Factual Earnings for Singles, Larger Effort Elasticity

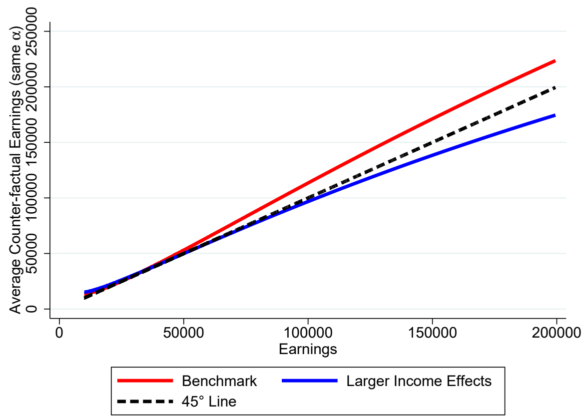
Note: Figures show average counter-factual earnings for each observed earnings level assuming all individuals had the same productivities (left) or the same preferences (right). The blue lines assumes elasticities are $\xi_z^c = \xi_z^u = 0.15$, $\xi_h^c = \xi_h^u = 0.05$ and the red lines assumes benchmark elasticities $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$.

C.3 Results for Single Individuals

We construct our counter-factual earnings measures using only single individuals (those with a household size of 1), thereby eliminating effects of dependents and spousal labor supply. Figures showing average counter-factual earnings by actual earnings level are plotted below, shown for all three of the elasticity estimates (benchmark, larger effort elasticity, and larger income effects) shown in the paper. The same general pattern holds for this restricted subset as in the main body using all earners.



(a) Homogeneous n



(b) Homogeneous α

Figure 10: Average Counter-Factual Earnings for Singles, Larger Income Effects

Note: Figures show average counter-factual earnings for each observed earnings level assuming all individuals had the same productivities (left) or the same preferences (right). The blue lines assumes elasticities are $\xi_z^c = \xi_h^c = 0.15$, $\xi_z^u = \xi_h^u = 0$ and the red lines assumes benchmark elasticities $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$.

D Optimal Tax Simulation Appendix

We start with a distribution of productivities and preferences $f(n, \alpha)$ computed using our method to recover individual (n, α) from labor supply elasticities as in Section V. We then choose a utility function that is consistent with these elasticities. In our benchmark elasticity case, we use:

$$U^{(1)}(c, e, h; n, \alpha) = \log\left(\alpha c - \frac{(eh)^{1+\frac{1}{0.15}}}{1 + \frac{1}{0.15}}\right)$$

which exhibits the constant labor supply elasticities $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$. Note all individuals are indifferent between any given value of eh. So that optimal choices of h^* and z^* are consistent with the assumed elasticities, we break this indifference using Equation 12 to determine $\log(e^*) = (\xi_z^c - \xi_h^c)/(\xi_h^c)\log(h^*)$, so that $e^*(n, \alpha) = 1\ \forall n, \alpha$. Next, we determine welfare weights for each (n, α) person under the preference neutrality assumption from Fleurbaey and Maniquet (2006). Specifically, preference neutrality implies that optimal tax rates are 0 if all inequality is due to preference heterogeneity. Operationally, 0 tax rates everywhere will be optimal if all individuals have the same marginal social value of consumption under 0 taxes (if the social marginal value of consumption is equal across individuals there is no motive to redistribute).[17] Normalizing weights $\mu(n, 1) = 1$, we get that $\mu(n, \alpha)U_c^{(1)}(c^*, e^*, h^*; n, \alpha) = U_c^{(1)}(c^*, e^*, h^*; n, 1)$ under 0

[17] Technically, this is only sufficient for a local optimal tax schedule - we assume it is also a global optima.

taxes. For our choice of utility function this implies (noting $c^* = nh^*e^*$):

$$\mu(n,\alpha)\frac{\alpha}{\alpha nh^*(n,\alpha)e^*(n,\alpha)-\frac{(e^*(n,\alpha)h^*(n,\alpha))^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}=\frac{1}{nh^*(n,1)e^*(n,1)-\frac{(e^*(n,1)h^*(n,1))^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}$$

Using the fact that $e^*(n,\alpha)h^*(n,\alpha) = (n\alpha)^{0.15}$ from the individual FOC:

$$\mu(n,\alpha)=\frac{\alpha n(n\alpha)^{0.15}-\frac{\left((n\alpha)^{0.15}\right)^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}{\alpha n(n)^{0.15}-\alpha\frac{(n^{0.15})^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}=\frac{\alpha n(n\alpha)^{0.15}-\alpha^{1+0.15}\frac{\left((n)^{0.15}\right)^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}{\alpha n(n)^{0.15}-\alpha\frac{(n^{0.15})^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}}=\alpha^{0.15}$$

The government's welfare function can be re-written as:

$$\max_{T(z)}\int_0^\infty\int_0^\infty \alpha^{0.15}\log\left(\alpha c^*(n,\alpha)-\frac{(e^*(n,\alpha)h^*(n,\alpha))^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}\right)f(n,\alpha)dnd\alpha$$

$$=\max_{T(z)}\int_0^\infty\int_0^\infty \alpha^{0.15}\log\left(\alpha\left(z^*(n,\alpha)-T(z^*(n,\alpha))-\frac{\left(\frac{z^*(n,\alpha)}{n\alpha^{0.15/1.15}}\right)^{\frac{1.15}{0.15}}}{\frac{1.15}{0.15}}\right)\right)f(n,\alpha)d\alpha dn$$

$$=\max_{T(z)}\int_0^\infty\int_0^\infty \alpha^{0.15}\log\left(z^*(n,\alpha)-T(z^*(n,\alpha))-\frac{\left(\frac{z^*(n,\alpha)}{n\alpha^{0.15/1.15}}\right)^{\frac{1.15}{0.15}}}{\frac{1.15}{0.15}}\right)+\alpha^{0.15}\log(\alpha)f(n,\alpha)d\alpha dn$$

$$=\max_{T(z)}\int_0^\infty\int_0^\infty \alpha^{0.15}\log\left(z^*(v)-T(z^*(v))-\frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1.15}{0.15}}}{\frac{1.15}{0.15}}\right)f(\alpha|v)d\alpha f(v)dv$$

$$=\max_{T(z)}\int_0^\infty \overline{\alpha^{0.15}}(v)\log\left(z^*(v)-T(z^*(v))-\frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1.15}{0.15}}}{\frac{1.15}{0.15}}\right)f(v)dv$$

The first equality swaps the integrals and uses $z^*(n,\alpha)/(n) = e^*(n,\alpha)h^*(n,\alpha)$ and $c^*(n,\alpha) = z^*(n,\alpha) - T(z^*(n,\alpha))$. The second equality is algebra. The third equality uses the fact that adding a constant $\alpha^{0.15}\log(\alpha)$ to the welfare function does not change the optimal tax schedule so can be safely ignored and does a change of variables from (n,α) to (v,α) (the Jacobian determinant is equal to 1). Following Lockwood and Weinzierl (2016), we refer to $v = n\alpha^{0.15/1.15}$ as the unified type. We can easily compute $f(\alpha|v)$ and $f(v)$ from $f(n,\alpha)$. The fourth equality evaluates the inner integral, denoted by $\overline{\alpha^{0.15}}(v)$, recognizing that

$$\log\left(z^*(v)-T(z^*(v))-\frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1.15}{0.15}}}{\frac{1.15}{0.15}}\right)$$

is not a function of α.

Finally, we have expressed the problem as a standard uni-dimensional optimal tax problem in terms of the unified type v. Hence, we can use the standard Hamiltonian optimization techniques to solve the problem as in Mirrlees (1971) or Saez (2001). I.e., the optimal tax rates are found by solving a system of ODE's derived from the envelope condition and the law of motion for the costate variable of the Hamiltonian (Equations (10) and (12) in Mirrlees (1971)).

For completeness, we show that we can also express the optimal tax problem as a uni-dimensional problem under the other two sets of elasticity parameters considered in Section V. First, if $\xi_z^c = \xi_z^u = 0.15, \xi_h^c = \xi_h^u = 0.05$, the utility function

$$U^{(1)}(c, e, h; n, \alpha) = \log\left(\alpha c - \frac{(eh)^{1+\frac{1}{0.15}}}{1+\frac{1}{0.15}}\right)$$

is still consistent with these elasticities. The only difference is that because all individuals are indifferent between any given value of eh, we must break this indifference by assuming $e^*(n, \alpha) = (\xi_z^c - \xi_h^c)/(\xi_h^c)h^*(n, \alpha) = 2h^*(n, \alpha)$. But other than that the problem is identical, hence standard uni-dimensional Hamiltonian optimization can still be used.

If $\xi_z^c = \xi_h^c = 0.15, \xi_z^u = \xi_h^u = 0$, then we use

$$U^{(2)}(c, e, h; n, \alpha) = \alpha \log(c) - \frac{(eh)^{\frac{1}{0.15}}}{\frac{1}{0.15}}$$

which exhibits the constant labor supply elasticities $\xi_z^c = \xi_h^c = 0.15, \xi_z^u = \xi_h^u = 0$ under constant taxes. Again, because hours and earnings elasticities are identical, we break individual indifference over he by assuming $e^*(n, \alpha) = 1$. Preference neutrality implies that welfare weights satisfy:

$$\mu(n, \alpha)\frac{\alpha}{c^*(n, \alpha)} = \frac{1}{c^*(n, 1)}$$

Under 0 taxes, $c^*(n, \alpha) = n\alpha^{0.15}$, hence:

$$\mu(n, \alpha) = \frac{n\alpha^{0.15}}{\alpha n} = \alpha^{0.15-1}$$

We can rewrite the optimal tax problem as follows:

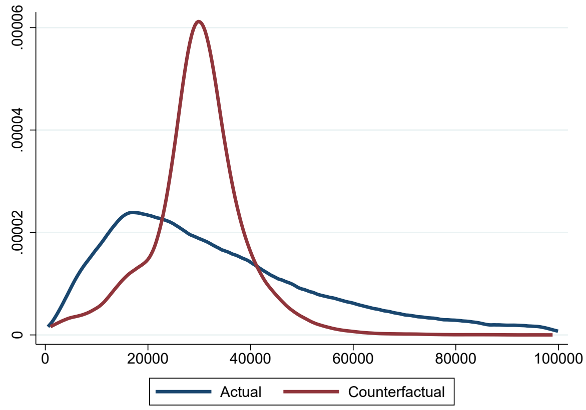
$$
\begin{aligned}
&\max_{T(z)} \int_0^\infty \int_0^\infty \alpha^{0.15-1} \left(\alpha \log(c^*(n,\alpha)) - \frac{(e^*(n,\alpha)h^*(n,\alpha))^{\frac{1}{0.15}}}{\frac{1}{0.15}} \right) f(n,\alpha) dn d\alpha \\
&= \max_{T(z)} \int_0^\infty \int_0^\infty \alpha^{0.15} \left(\log(z^*(n,\alpha) - T(z^*(n,\alpha)) - \frac{\left(\frac{z^*(n,\alpha)}{n\alpha^{0.15}}\right)^{\frac{1}{0.15}}}{\frac{1}{0.15}} \right) f(n,\alpha) d\alpha dn \\
&= \max_{T(z)} \int_0^\infty \int_0^\infty \alpha^{0.15} \left(\log(z^*(v) - T(z^*(v)) - \frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1}{0.15}}}{\frac{1}{0.15}} \right) f(\alpha|v) d\alpha f(v) dv \\
&= \max_{T(z)} \int_0^\infty \overline{\alpha^{0.15}}(v) \left(\log(z^*(v) - T(z^*(v)) - \frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1}{0.15}}}{\frac{1}{0.15}} \right) f(v) dv
\end{aligned}
$$

The first equality swaps the integrals, multiplies and divides by α and uses $(z^*(n,\alpha))/(n) = e^*(n,\alpha)h^*(n,\alpha)$ and $c^*(n,\alpha) = z^*(n,\alpha) - T(z^*(n,\alpha))$. The second equality does a change of variables from (n,α) to (v,α) (the Jacobian determinant is equal to 1). The unified type is denoted $v = n\alpha^{0.15}$. We can again easily compute $f(\alpha|v)$ and $f(v)$ from $f(n,\alpha)$. The third equality evaluates the inner integral, denoted by $\overline{\alpha^{0.15}}(v)$, recognizing that

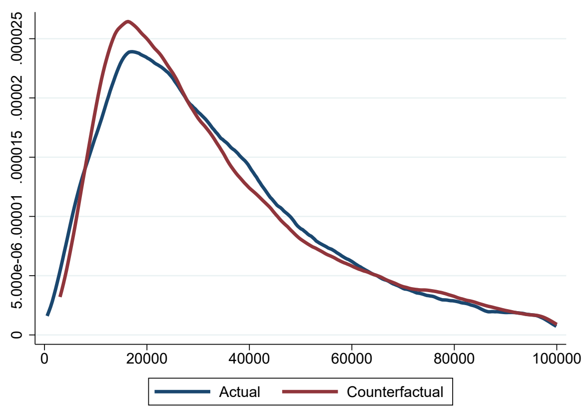
$$
\left(\log(z^*(v) - T(z^*(v)) - \frac{\left(\frac{z^*(v)}{v}\right)^{\frac{1}{0.15}}}{\frac{1}{0.15}} \right)
$$

is not a function of α. Again, this last optimization problem is a standard one dimensional tax problem in terms of the unified type v so we can use Hamiltonian techniques to solve for the optimal rates.

E Miscellaneous Figures Appendix



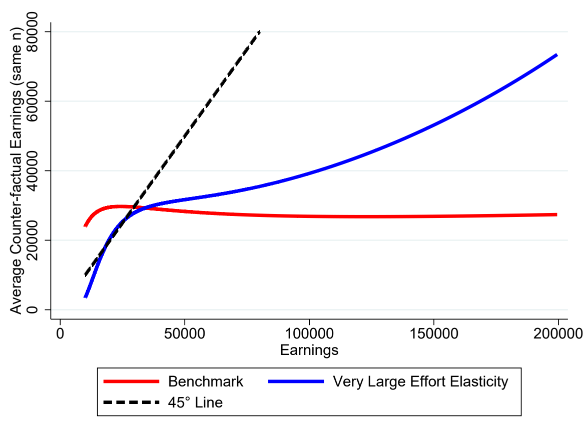
(a) Homogeneous n



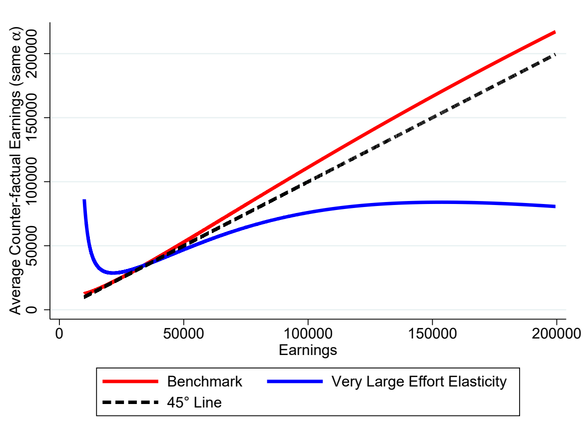
(b) Homogeneous α

Figure 11: Counter-Factual Earnings Distribution, Benchmark Elasticities

Note: Elasticities used to construct this figure are from Chetty (2012): $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$.



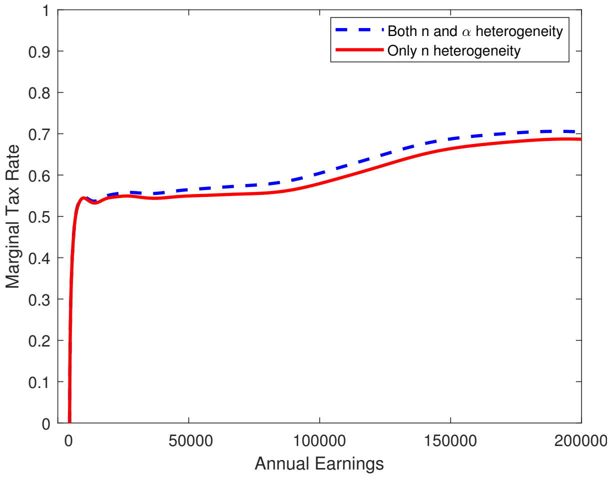
(a) Homogeneous n



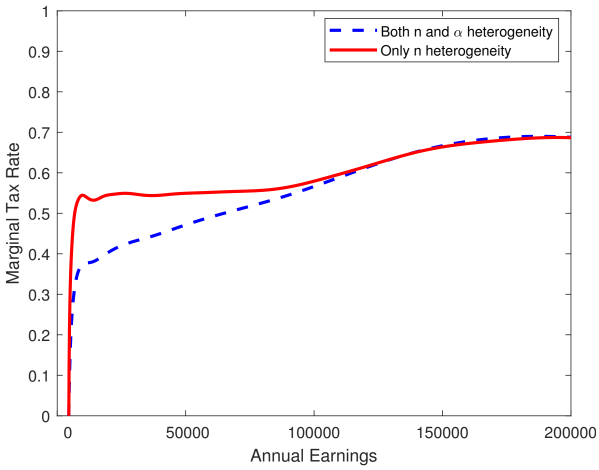
(b) Homogeneous α

Figure 12: Average Counter-Factual Earnings, Even Larger Effort Elasticity

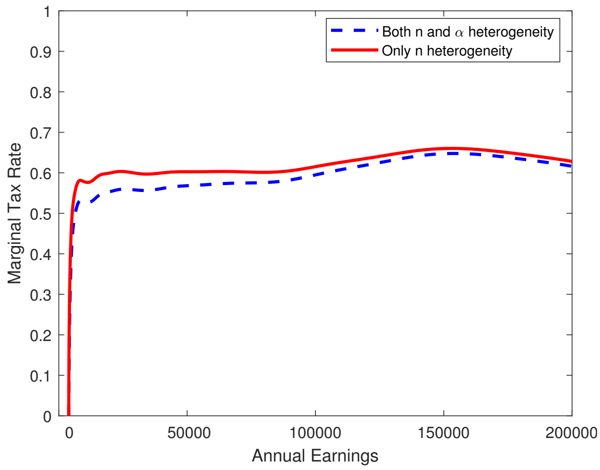
Note: Figures show average counter-factual earnings for each observed earnings level assuming all individuals had the same productivities (left) or the same preferences (right). The blue lines assume elasticities are $\xi_z^u = \xi_z^c = 0.15$, $\xi_h^u = \xi_h^c = 0.025$ and the red lines assume benchmark elasticities $\xi_z^c = \xi_z^u = \xi_h^c = \xi_h^u = 0.15$.



(a) Benchmark Case



(b) Higher Effort Elasticity



(c) Large Income Effects

Figure 13: Optimal Marginal Tax Rates with Productivity and Preference Heterogeneity

Note: Elasticities in panel (a) are $\xi_z^c = \xi_h^c = \xi_z^u = \xi_h^u = 0.15$. Elasticities in panel (b) are $\xi_z^c = \xi_z^u = 0.15, \xi_h^c = \xi_h^u = 0.05$. Elasticities in panel (c) are $\xi_z^c = \xi_h^c = 0.15, \xi_z^u = \xi_h^u = 0$.

References

Bureau of Labor Statistics. 2016. "American Time Use Survey [dataset]." *United States Department of Labor. https://www.bls.gov/tus/datafiles-0319.htm (accessed November 3, 2016).*

Bureau of Labor Statistics. 2020. "CPI for All Urban Consumers (CPI-U), Seasonally adjusted - CUSR0000SA0." *United States Department of Labor. https://data.bls.gov/timeseries/CUSR0000SA0?output_view=pct_1mth (accessed July 24, 2020).*

Chetty, Raj. 2012. "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica*, 80(3): 969–1018.

Families and Work Institute. 2008. "National Study of the Changing Workforce [dataset]." *https://searchworks.stanford.edu/view/10322573 (accessed May 30, 2018).*

Fleurbaey, Marc, and Francois Maniquet. 2006. "Fair Income Tax." *Review of Economic Studies*, 73: 55–83.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2018. "Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]." *Minneapolis, MN: IPUMS. https://doi.org/10.18128/D030.V7.0 (accessed September 5, 2018).*

Gale, David, and Hukukane Nikaido. 1965. "The Jacobian Matrix and Global Univalence of Mappings." *Math. Annalen*, 159: 81–93.

Heim, Bradley T. 2010. "The responsiveness of self-employment income to tax rate changes." *Labour Economics*, 17: 940–950.

Lockwood, Benjamin B., and Matthew Weinzierl. 2016. "De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution." *Journal of Public Economics*, 124: 74–80.

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

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