Online Appendix: Not for Publication

A Derivation of Willingness to Pay

I first characterize $\frac{\frac{d\hat{V}_i}{d\theta}|_{\theta=0}}{\lambda_i}$. Taking the total derivative of V_i with respect to θ, I have

$$\frac{d\hat{V}_i}{d\theta} = \frac{dV_i\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{l}}, \hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{T}_i, y_i, \hat{G}_i\right)}{d\theta} = \frac{\partial V_i}{\partial T_i}\frac{d\hat{T}_i}{d\theta} + \sum_{j=1}^{J_G}\frac{\partial V_i}{\partial G_i}\frac{d\hat{G}_i}{d\theta} + \sum_{j=1}^{J_X}\frac{\partial V_i}{\partial \tau_{ij}^x}\frac{d\hat{\tau}_{ij}^x}{d\theta} + \sum_{j=1}^{J_L}\frac{\partial V_i}{\partial \tau_{ij}^l}\frac{d\hat{\tau}_{ij}^l}{d\theta}$$

Applying the envelope theorem from the agent's maximization problem and evaluating at $\theta = 0$ implies

$$\begin{aligned}
\frac{\partial V_i}{\partial \tau_{ij}^x} &= -x_{ij}\lambda_i \\
\frac{\partial V_i}{\partial \tau_{ij}^l} &= -l_{ij}\lambda_i \\
\frac{\partial V_i}{\partial T_i} &= -\lambda_i \\
\frac{\partial V_i}{\partial G_i} &= \frac{\partial u_i}{\partial G_i}
\end{aligned}$$

Replacing terms, I have

$$\frac{d\hat{V}_i}{d\theta}|_{\theta=0} = \lambda_i\left(\frac{d\hat{T}_i}{d\theta} + \sum_{j=1}^{J_G}\frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i}\frac{d\hat{G}_{ij}}{d\theta} - \sum_{j=1}^{J_X}x_{ij}\frac{d\hat{\tau}_{ij}^x}{d\theta} - \sum_{j=1}^{J_L}l_{ij}\frac{d\hat{\tau}_{ij}^l}{d\theta}\right)$$

Now, I use equation 5 to replace the total transfers, $\frac{d\hat{T}_i}{d\theta}$, with the net government budgetary position, $\frac{d\hat{t}_i}{d\theta}$, which yields

$$\frac{d\hat{V}_i}{d\theta}|_{\theta=0} = \lambda_i\left(\sum_{j=1}^{J_G}\left(\frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i} - c_j^G\right)\frac{d\hat{G}_{ij}}{d\theta} + \frac{d\hat{t}_i}{d\theta} + \frac{d}{d\theta}\left[R\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{\mathbf{x}}_{\mathbf{i}}, \hat{\tau}_{\mathbf{i}}^{\mathbf{l}}, \hat{\mathbf{l}}_{\mathbf{i}}\right)\right] - \sum_{j=1}^{J_X}x_{ij}\frac{d\hat{\tau}_{ij}^x}{d\theta} - \sum_{j=1}^{J_L}l_{ij}\frac{d\hat{\tau}_{ij}^l}{d\theta}\right)$$

Finally, note that equation 6 shows I can replace the difference between the total revenue impact, $\frac{d}{d\theta}\left[R\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{\mathbf{x}}_{\mathbf{i}}, \tau_{\mathbf{i}}^{\mathbf{l}}, \hat{\mathbf{l}}_{\mathbf{i}}\right)\right]$, and the mechanical revenue effect, $\sum_{j=1}^{J_X}x_{ij}\frac{d\hat{\tau}_{ij}^x}{d\theta} + \sum_{j=1}^{J_L}l_{ij}\frac{d\hat{\tau}_{ij}^l}{d\theta}$, with the behavioral impact of the policy on the government budget constraint, yielding

$$\frac{dV_i}{d\theta}|_{\theta=0} = \lambda_i\left(\sum_{j=1}^{J_G}\left(\frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i} - c_j^G\right)\frac{d\hat{G}_{ij}}{d\theta} + \frac{d\hat{t}_i}{d\theta} + \left(\sum_j \tau_{ij}^x\frac{d\hat{x}_{ij}}{d\theta} + \sum_j \tau_{ij}^l\frac{d\hat{l}_{ij}}{d\theta}\right)\right)$$

B Non-Marginal Welfare Analysis

In reality, policy changes are not always small. In such cases, one might be worried that the use of the difference in potential outcomes may not reflect a local derivative, $\frac{dx}{d\theta}$. Here, I provide conditions under which one can use the difference in causal effects to construct a measure of the (non-marginal) equivalent variation of the policy change. Heuristically, one can use the framework to estimate equivalent variation as long as the policy does not induce a significant effect on the marginal utility of income.

Equivalent variation, $EV(\theta)$, of the policy at point θ from the initial point $\theta = 0$ is given by the implicit equation:

$$V(P, y + EV(\theta)) = \hat{V}(\theta)$$

where $V(P, y)$ is the utility obtained under policy P with non-labor income $y + EV(\theta)$. Differentiating yields:

$$EV'(\theta) = \frac{\frac{d\hat{V}(\theta)}{d\theta}}{\lambda(P, y + EV(\theta))} = \underbrace{\frac{\lambda\left(\hat{P}(\theta), y\right)}{\lambda(P, y + EV(\theta))}}_{\text{MU of Income Adjustment}} \underbrace{\frac{\frac{d\hat{V}}{d\theta}}{\lambda\left(\hat{P}(\theta), y\right)}}_{\text{Std Measure}}$$

where $\frac{\frac{d\hat{V}}{d\theta}}{\lambda\left(\hat{P}(\theta),y\right)}$ relies on the local causal effects of the policy at $P(\theta)$. Expanding yields:

$$EV(1) = \int_0^1 \frac{\lambda\left(\hat{P}(\theta), y\right)}{\lambda(P, y + EV(\theta))} \left[\left(\frac{\frac{\partial \hat{u}}{\partial \hat{G}}}{\lambda\left(\hat{P}(\theta), y\right)} - c_G\right)\frac{d\hat{G}}{d\theta} + \frac{d\hat{t}}{d\theta} + \sum_j \hat{\tau}_j \frac{d\hat{x}_j}{d\theta}\right] d\theta$$

Conditions for Global = Local If two conditions are satisfied, global and local conditions are equivalent. Suppose that:

(a) the marginal utility of income does not vary for the policy relative to the income effects: $\lambda\left(\hat{P}(\theta), y\right) = \lambda(P, y + EV(\theta))$

(b) the causal effects are linear in θ (i.e. $\frac{d\hat{x}_j}{d\theta} = \hat{x}_j(1) - \hat{x}_j(0)$ and $\frac{d\hat{l}_j}{d\theta} = \hat{l}_j(1) - \hat{l}_j(0)$ for all θ.

Note that (a) is implied by quasilinear utility, but is far less restrictive. Also, (b) is commonly imposed in empirical applications. To derive the total equivalent variation for the policy, let $D_j = \int_0^1 \left(\frac{\frac{\partial \hat{u}}{\partial G_j}}{\lambda\left(\hat{P}(\theta),y\right)} - c_j^G\right) d\theta$ denote the average willingness to pay above cost for the publicly provided goods. Then, if (a) and (b) hold, one can show that:

$$EV(1) = \underbrace{\sum_j \Delta\hat{G}_j * D_j +}_{\text{Public Goods}} \underbrace{\Delta\hat{t}}_{\text{Net Transfer}} + \underbrace{\sum_j \bar{\tau}_j^x \Delta\hat{x}_j + \sum_j \bar{\tau}_j^l \Delta\hat{l}_j}_{\text{Behavioral Reponse}}$$

where $\Delta\hat{G}_j = \hat{G}_j(1) - \hat{G}_j(0)$ is the change in publicly provided good j, $\Delta\hat{t}$is the change in net resources, and $\Delta\hat{x}_j = \hat{x}_j(1) - \hat{x}_j(0)$ is the difference in potential outcomes in policy world $\theta = 1$ relative to $\theta = 0$

(i.e. $\Delta \hat{x}_j$ is the **non-marginal** causal effect of the policy on $\hat{x}_j$).

C Appendix: Externalities (and Internalities)

The fact that the causal effect does not need to be decomposed into income and substitution effects extends to a more complex environment with internalities and externalities.

To see this, now suppose that the agents' utility function is given by

$$u_i\left(\mathbf{x_i}, \mathbf{l_i}, \mathbf{G_i}, E_i\right)$$

where the externality imposed on agent i, E_i, is produced in response to the consumption choices of all agents in the economy,

$$E_i = f_i^E\left(\mathbf{x}\right)$$

where $\mathbf{x} = \{\mathbf{x_i}\}_{\mathbf{i}}$ is the consumption decisions made by the agent (one could generalize this easily to incorporate l). I assume that there is no market for E_i and that agents do not take E_i into account when conducting their optimization. Note that I allow E_i to interact arbitrarily with the utility function, but I assume it is taken as given in the agents' maximization problem. Thus, E_i could represent a classical externality (e.g. pollution) or a behavioral "internality". An internality could be welfare costs of smoking that are not incorporated into their maximization program, or could incorporate "optimization frictions" of the form used by Chetty (2009a) where taxpayers over-estimate the costs of tax sheltering so that the marginal utility of tax sheltered income is not equal to the marginal utility of taxable income.

The value function is now given by

$$\begin{aligned} V_i\left(\tau_{\mathbf{i}}^{\mathbf{l}}, \tau_{\mathbf{i}}^{\mathbf{x}}, T_i, y_i, \mathbf{G_i}, E_i\right) &= \max_{\mathbf{x},\mathbf{l}} u_i\left(\mathbf{x}, \mathbf{l}, \mathbf{G_i}, E_i\right) \\ & s.t. \quad \sum_{j=1}^{J_X}\left(1+\tau_{ij}^x\right)x_{ij} \leq \sum_{j=1}^{J_L}\left(1-\tau_{ij}^l\right)l_{ij} + T_i + y_i \end{aligned}$$

Given each agent's solution to this program, $\mathbf{x_i}$, I construct $E_i = f_i^E\left(\mathbf{x}\right)$ and $\mathbf{x}$ is the vector of solutions to each agents optimization program.

All other definitions from Section 2 are maintained. In particular, policy paths are defined as in equation 4.[51] Proposition 2 presents the characterization of the marginal welfare impact of a policy evaluated at $\theta = 0$.

[51] Note that I do not allow the government to directly affect the level of E. This would be duplicating the role of publicly provided goods, as I could specify G to be provision of goods which mitigate the externality (either directly or through their effect on agents' choices of x).

Proposition 1. *The welfare impact of the marginal policy change to type i is given by*

$$\frac{\frac{dV_i}{d\theta}|_{\theta=0}}{\lambda_i} = \left(\underbrace{\sum_{j=1}^{J_G} \left(\frac{\frac{\partial u_i}{\partial G}}{\lambda_i} - c_j^G \right) \frac{d\hat{G}_{ij}}{d\theta}}_{\substack{\textit{Public Spending /} \\ \textit{Market Failure}}} + \underbrace{\frac{d\hat{t}_i}{d\theta}}_{\textit{Transfer}} + \underbrace{\left(\sum_j^{J_X} \tau_{ij}^x \frac{d\hat{x}_{ij}}{d\theta} + \sum_j^{J_L} \tau_{ij}^l \frac{d\hat{l}_{ij}}{d\theta} \right)}_{\textit{Impact on Govt}} + \underbrace{\frac{\frac{\partial u_i}{\partial E_i}}{\lambda_i} \frac{d\hat{E}_i}{d\theta}}_{\textit{Impact on Externality}} \right)$$

where

$$\frac{d\hat{E}_i}{d\theta} = \left(\sum_i \sum_j^{J_X} \frac{\partial f_i^E}{\partial x_{ij}} \frac{d\hat{x}_{ij}}{d\theta} \right)$$

is the net marginal impact of the policy on the externality experienced by type i.

Proof. Taking the total derivative of V_i with respect to θ, I have

$$\frac{dV_i\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{l}}, \hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{T}_i, y_i, \hat{\mathbf{G}}_{\mathbf{i}}, \hat{E}_i\right)}{d\theta} = \frac{\partial V_i}{\partial T_i}\frac{d\hat{T}_i}{d\theta} + \sum_{j=1}^{J_G} \frac{\partial V_i}{\partial G_{ij}} \frac{d\hat{G}_{ij}}{d\theta} + \sum_{j=1}^{J_X} \frac{\partial V_i}{\partial \tau_{ij}^x} \frac{d\hat{\tau}_{ij}^x}{d\theta} + \sum_{j=1}^{J_L} \frac{\partial V_i}{\partial \tau_{ij}^l} \frac{d\hat{\tau}_{ij}^l}{d\theta} + \frac{\partial V_i}{\partial E_i}\frac{d\hat{E}_i}{d\theta}$$

Applying the envelope theorem from the agent's maximization problem and evaluating at $\theta = 0$ implies

$$\begin{aligned} \frac{\partial V_i}{\partial \tau_{ij}^x} &= -x_{ij}\lambda_i \\ \frac{\partial V_i}{\partial \tau_{ij}^l} &= -l_{ij}\lambda_i \\ \frac{\partial V_i}{\partial T_i} &= -\lambda_i \\ \frac{\partial V_i}{\partial G_{ij}} &= \frac{\partial u_i}{\partial G_{ij}} \\ \frac{\partial V_i}{\partial E_i} &= \frac{\partial u_i}{\partial E_i} \end{aligned}$$

Replacing terms, I have

$$\frac{dV_i}{d\theta}|_{\theta=0} = \lambda_i \left(\frac{d\hat{T}_i}{d\theta} + \sum_{j=1}^{J_G} \frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i} \frac{d\hat{G}_{ij}}{d\theta} - \sum_{j=1}^{J_X} x_{ij} \frac{d\hat{\tau}_{ij}^x}{d\theta} - \sum_{j=1}^{J_L} l_{ij} \frac{d\hat{\tau}_{ij}^l}{d\theta} + \frac{\partial u_i}{\partial E_i}\frac{d\hat{E}_i}{d\theta} \right)$$

Now, I use equation 5 to replace the total transfers, $\frac{d\hat{T}_i}{d\theta}$, with the net government budgetary position,

$\frac{d\hat{t}_i}{d\theta}$, which yields

$$\frac{dV_i}{d\theta}|_{\theta=0} = \lambda_i \left(\sum_{j=1}^{J_G} \left(\frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i} - c_j^G \right) \frac{d\hat{G}_{ij}}{d\theta} + \frac{d\hat{t}_i}{d\theta} + \frac{d}{d\theta} \left[R\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{\mathbf{x}}_{\mathbf{i}}, \hat{\tau}_{\mathbf{i}}^{\mathbf{l}}, \hat{\mathbf{l}}_{\mathbf{i}} \right) \right] - \sum_{j=1}^{J_X} x_{ij} \frac{d\hat{\tau}_{ij}^x}{d\theta} - \sum_{j=1}^{J_L} l_{ij} \frac{d\hat{\tau}_{ij}^l}{d\theta} + \frac{\partial u_i}{\partial E_i} \frac{d\hat{E}_i}{d\theta} \right)$$

Finally, note that equation 6 shows I can replace the difference between the total revenue impact, $\frac{d}{d\theta} \left[R\left(\hat{\tau}_{\mathbf{i}}^{\mathbf{x}}, \hat{\mathbf{x}}_{\mathbf{i}}, \tau_{\mathbf{i}}^{\mathbf{l}}, \hat{\mathbf{l}}_{\mathbf{i}} \right) \right]$, and the mechanical revenue effect, $\sum_{j=1}^{J_X} x_{ij} \frac{d\hat{\tau}_{ij}^x}{d\theta} + \sum_{j=1}^{J_L} l_{ij} \frac{d\hat{\tau}_{ij}^l}{d\theta}$, with the behavioral impact of the policy on the government budget constraint, yielding

$$\frac{dV_i}{d\theta}|_{\theta=0} = \lambda_i \left(\sum_{j=1}^{J_G} \left(\frac{\frac{\partial u_i}{\partial G_{ij}}}{\lambda_i} - c_j^G \right) \frac{d\hat{G}_{ij}}{d\theta} + \frac{d\hat{t}_i}{d\theta} + \left(\sum_{j}^{J_X} \tau_{ij}^x \frac{d\hat{x}_{ij}}{d\theta} + \sum_{j}^{J_L} \tau_{ij}^l \frac{d\hat{l}_{ij}}{d\theta} \right) + \frac{\partial u_i}{\partial E_i} \frac{d\hat{E}_i}{d\theta} \right)$$

And, note that I can expand $\frac{d\hat{E}_i}{d\theta}$ by taking a total derivative of $E_i = f_i^E(\mathbf{x})$ across all goods and types, yielding

$$\frac{d\hat{E}_i}{d\theta} = \sum_i \sum_{j=1}^{J_X} \frac{\partial f_i^E}{\partial x_{ij}} \frac{d\hat{x}_{ij}}{d\theta}$$

which concludes the proof. □

With externalities, I must know the net causal effect of behavioral response to the policy on the externality, $\frac{dE_i}{d\theta} = \left(\sum_j^{J_X} \frac{\partial f_i^E}{\partial x_{ij}} \frac{d\hat{x}_{ij}}{d\theta} \right)$, along with the the marginal willingness to pay for the externality, $\frac{\frac{\partial u_i}{\partial E_i}}{\lambda_i}$. Therefore, the welfare loss from a behavioral response that reduces government revenue may be counteracted by the welfare gain from any reduction on the externality imposed on other individuals. Thus, financing government revenue using so-called "green taxes" that also reduce externalities may deliver higher government welfare than policies whose financing schemes do not reduce externalities.[52] This is the so-called "double-dividend" highlighted in previous literature (Bovenberg and de Mooji (1994); Goulder (1995); Parry (1995)). But even in this world, the causal effect of the policy on behavior, i.e. the policy elasticity, continue to be the behavioral elasticities that are relevant for estimating welfare impact of the policy.

D General Equilibrium Effects

By assuming one unit of goods are produced with one unit of labor supply, the model ruled out general equilibrium effects (i.e. that the policy change affects prices). However, such effects are easily incorporated into the model by adding the implied transfers to the net resources term, $\frac{d\hat{t}_i}{d\theta}$.

For example, if the policy increases the price of i's labor supply activity j, then she will obtain a resource benefit of $l_{ij} \frac{dw_{ij}}{d\theta}|_{\theta=0}$, where $\frac{dw_{ij}}{d\theta}|_{\theta=0}$ is the causal impact of the policy on the after-tax wage faced by individual i on her jth labor supply activity. These additional impacts are valued

[52] As is well-known (e.g. Salanie (2003)), if taxes are initially near their optimal levels, then at the margin it is not clear that an additional green tax will be any more desirable than a tax on any other good.

dollar-for-dollar and can simply be added to the resource transfer term, $\frac{d\hat{t}_i}{d\theta}|_{\theta=0}$. Hence, when policies have general equilibrium effects, one also needs to track the causal impact of the policy on prices, and adjust the size of the transfers, $\frac{d\hat{t}_i}{d\theta}$ accordingly. The causal effects are still the desired responses, but one needs to also know the general equilibrium effects of government policies.

E Optimal Commodity Taxation and the "Inverse Elasticity" Rule

Ramsey (1927) proposes the question of how commodities should be taxed in order to raise a fixed government expenditure, $R > 0$. Diamond and Mirrlees (1971) provide a formal modeling of this environment and show that, at the optimum, the tax-weighted Hicksian price derivatives for each good are equated. Here, I illustrate this result and relate it to the framework provided in this paper.

Assume there is a representative agent and drop i subscripts. A necessary conditions for tax policy to be at an optimum is given by

$$\frac{d\hat{V}_P}{d\theta} = 0$$

for all feasible policy paths, P. With a representative agent, the optimal tax would be lump-sum of size R. However, the optimal commodity tax program proposed by Ramsey (1927) makes the assumption that the government cannot conduct lump-sum taxation. Hence, the only feasible policies are those that raise and lower tax rates in a manner that preserves the budget constraint.

Consider a policy, $P(\theta)$, that lowers the tax on good 1 and raises the tax on good 2. The optimality condition is given by

$$\sum_k \hat{\tau}_k \frac{d\hat{x}_k}{d\theta} = 0 \tag{18}$$

Equation (18) suggests more responsive goods should be taxed at lower rates, thereby nesting the standard "inverse elasticity" argument (higher $\frac{d\hat{x}_k}{d\theta}$ should be associated with lower $\hat{\tau}_k$). The optimal tax attempts to replicate lump-sum taxes by taxing relatively inelastic goods.

Diamond and Mirrlees (1971) further note that, because $\frac{d\hat{V}_P}{d\theta} = 0$ at the optimum, one can expand the behavioral change using the Hicksian demands, x_k^h,

$$\frac{dx_k}{d\theta} = \frac{\partial x_k^h}{\partial \tau_1}\frac{d\tau_1}{d\theta} + \frac{\partial x_k^h}{\partial \tau_2}\frac{d\tau_2}{d\theta}$$

where, in general, there would be the additional term, $\frac{\partial x_k^h}{\partial u}\frac{dV_p}{d\theta}$, but this vanishes at the optimum. Hence, that the optimality condition is given by

$$\sum_k \tau_k \frac{\partial x_k^h}{\partial \tau_1}\frac{d\tau_1}{d\theta} = \sum_k \tau_k \frac{\partial x_k^h}{\partial \tau_2}\left(-\frac{d\tau_2}{d\theta}\right) \tag{19}$$

so that the tax-weighted Hicksian responses are equated across the tax rates – precisely the classic result in Diamond and Mirrlees (1971) (see equation 38).[53]

[53]Under the additional assumption that compensated cross-price elasticities are zero, one arrives at the classic inverse

However, note that one never relied on compensated elasticities to test the optimality condition in equation (18). Compensated elasticities arise only because of the assumption that policy is at the optimum. One could consider any budget-neutral policy that simultaneously adjusts two commodity taxes and test equation (18) directly. Conditional on knowing the causal effects of such a policy, one would not need to know whether income or substitution effects drive the behavioral response to commodity taxes. The policy elasticities would be sufficient.

F Application Details

F.1 EITC

This section outlines the welfare analysis of an EITC expansion. To correspond with the causal effects analyzed in much previous literature, the marginal expansion of the EITC program can best be thought of as increasing the maximum benefit level in a manner that maintains current income eligibility thresholds and tax schedule kink points (but raises the phase-in and phase-out rates in order to reach the new maximum benefit). However, the results from Chetty et al. (2013) suggest the phase-out slope of the EITC has only a minor impact on labor supply (most of the response is from individuals below the EITC maximum benefit level choosing to increase their labor supply). This suggests the impact on the behavioral response on the government budget would not be too sensitive to the precise design of the phase-out of the program.

The effects documented in previous literature consist of both intensive and extensive labor supply responses. With extensive margin responses, $\frac{d\hat{l}_i^{EITC}}{d\theta}$ may not exist for all i, as individuals make discrete jumps in their choice of labor supply. However, this is easily accommodated into the model. To see this, normalize the index of the Poor to be the unit interval, $i \in Poor = [0,1]$. Then, order the index of the poor population such that $\hat{l}_i(\theta) > 0$ implies $\hat{l}_j(\theta) > 0$ for $j < i$ and all $\theta \in (-\epsilon, \epsilon)$. With this ordering, there exists a threshold, $i^{LFP}(\theta)$, such that $i < i^{LFP}(\theta)$ indicates that i is in the labor force and $i > i^{LFP}(\theta)$ indicates that i is not in the labor force. Hence, $i^{LFP}(\theta)$ is the fraction of the poor single mothers that are in the labor force. With this notation, the impact of the behavioral response to the policy by the poor on the government's budget is given by:

$$-\int_{i\in Poor} \tau_i^l \frac{d\hat{l}_i^{EITC}}{d\theta}|_{\theta=0} di = \underbrace{\left(\tau^l_{i^{LFP}(0)} l_{i^{LFP}(0)}\right) \frac{di^{LFP}}{d\theta}|_{\theta=0}}_{\text{Extensive Margin}} - \underbrace{\int_{i<\hat{i}} \tau_i^l \frac{d\hat{l}_i^{EITC}}{d\theta}|_{\theta=0} di}_{\text{Intensive Margin}} \quad (20)$$

where $\tau^l_{i^{LFP}(0)} l_{i^{LFP}(0)}$ is the average taxable income (or loss) generated by the marginal type entering the labor force and $\frac{di^{LFP}}{d\theta}$ is the marginal rate at which the policy induces labor force entry. The cost

elasticity rule:

$$\frac{\tau_2}{\tau_1} = \frac{\frac{\partial x_1^h}{\partial \tau_1}\frac{d\tau_1}{d\theta}}{\frac{\partial x_2^h}{\partial \tau_2}\frac{d\tau_2}{d\theta}}$$

so that optimal tax rates are inversely proportional to their compensated (Hicksian) demands.

resulting from extensive margin responses is given by the impact of the program on the labor force participation rate, multiplied by the size of the average subsidy to those entering the labor force.[54]

There is a large literature analyzing the impact of the EITC expansion on labor force participation of single mothers, beginning with Eissa and Liebman (1996). These approaches generally estimate the causal effect of EITC receipt on behavior using various expansions in the generosity of the EITC program. Hotz and Scholz (2003) summarize this literature and find consistency across methodologies in estimates of the elasticity of the labor force participation rate of single mothers, $\hat{\imath}$, rate with respect to the average after-tax wage, $E\left[\left(1-\tau^l_{\hat{i}}\right) l_{\hat{i}}\right]$, with estimates ranging from 0.69-1.16.

I translate this elasticity into equation (20) by normalizing θ to parameterize an additional unit of the mechanical subsidy[55] and writing:

$$\left(\tau^l_{i^{LFP}(0)} l_{i^{LFP}(0)}\right) \frac{di^{LFP}}{d\theta}|_{\theta=0} = \frac{\left(\tau^l_{i^{LFP}(0)} l_{i^{LFP}(0)}\right)}{\left(\left(1-\tau^l_{i^{LFP}(0)}\right) l_{i^{LFP}(0)}\right)} \epsilon^{LFP}_{E\left[\left(1-\tau^l_{\hat{i}}\right) l_{\hat{i}}\right]}$$

where $\epsilon^{LFP}_{E\left[\left(1-\tau^l_{\hat{i}}\right) l_{\hat{i}}\right]}$ is the elasticity of the labor force participation rate with respect to the after tax wage rate and $\frac{E\left[\tau^l_{\hat{i}} l_{\hat{i}}\right]}{E\left[\left(1-\tau^l_{\hat{i}}\right) l_{\hat{i}}\right]}$ is the size of the subsidy as a fraction of after tax income for the marginal labor force entrant. For the elasticity of labor force participation, I choose an estimate of 0.9, equal to the midpoint of existing estimates (Hotz and Scholz (2003)). For $\frac{E\left[\tau^l_{\hat{i}} l_{\hat{i}}\right]}{E\left[\left(1-\tau^l_{\hat{i}}\right) l_{\hat{i}}\right]}$, one desires the after tax wages and subsidies for marginal entrants into the labor force. While such parameters could be identified using the same identification strategies previous papers have used to estimate the labor supply impact of the EITC, to my knowledge no such estimates of the marginal wages and subsidies exist. Using the 2004 SOI, Eissa and Hoynes (2011) report that the average subsidy is $1,806 per beneficiary, which corresponds to 9.2% of a $20,000 gross income for EITC beneficiaries. Athreya et al. (2010) report the average recipient obtains a subsidy equal to 11.7% of gross income in the 2008 CPS. I therefore take the approximate midpoint of 11%.

These calculations suggest the extensive margin impact on the government budget is given by:

$$\underbrace{E\left[\tau^l_{\hat{i}} l_{\hat{i}}\right] \frac{d\hat{i}}{d\theta}}_{\text{Extensive Margin}} = \frac{0.11}{1+0.11} * 0.9 = 0.09$$

so that the EITC is 9% more costly to the government because of extensive margin labor supply responses.[56]

[54] Because the model assumed individuals face linear tax rates, the distinction between the average and marginal tax rate is not readily provided, but it is straightforward to verify that the fiscal externality imposed by those entering the labor force is given by the size of the subsidy they receive by entering the labor force, not by the marginal tax or subsidy they face if they were to provide an additional unit of labor supply.

[55] This normalizes $\int_{i \in Poor} \left(\frac{d\hat{T}^{EITC}_i}{d\theta}|_{\theta=0} + \frac{d\hat{\tau}^{EITC}_{ij}}{d\theta}|_{\theta=0} l_i\right) di = 1$

[56] Taking elasticity estimates in the 0.69-1.12 range reported by Hotz and Scholz (2003), yields estimates of the extensive margin impact ranging from 0.07 to 0.11. Hence, if one assumed only extensive margin responses were operating, the

Intensive margin responses Until recently, there was little evidence that the EITC had intensive margin impacts on labor supply. However, the recent paper by Chetty et al. (2013) exploits the geographic variation in knowledge about the marginal incentives induced by the EITC, as proxied by the local fraction of self-employed that bunch at the subsidy-maximizing kink rate. Using the universe of tax return data from EITC recipients, their estimates suggest that the behavioral responses induced by knowledge about the marginal incentives provided by the EITC increase refunds by approximately 5% relative to what they would be in the absence of behavioral responses, with most of these responses due to intensive margin adjustments. What is particularly useful about this study is that it uses tax expenditures as an outcome variable, and hence can compute the associated fiscal externality directly.

The downside of Chetty et al. (2013) is that the policy path in question is the degree of "knowledge about the shape of the EITC schedule". While this policy path provides guidance on the size of the distortions induced by these marginal incentives, one could imagine that even in places with no knowledge of the EITC schedules the existence of the EITC generates extensive margin responses.

To account for this, I make the baseline assumption that the knowledge of the average EITC subsidy generates extensive margin responses and knowledge of the shape of the EITC schedule generates intensive margin responses. With this assumption, the results of Chetty et al. (2013) should be added together with the extensive margin responses found in previous literature to arrive at the total impact of an EITC expansion. This yields an estimate of $FE^{EITC} = 0.09 + 0.05 = 14\%$ with a range of 0.12-0.16 taking the range of extensive margin labor supply responses.[57]

F.2 Food Stamps

Using variation induced in the introduction of food stamps in the 1960s and 70s Hoynes and Schanzenbach (2012) estimate that food stamps led to a significant reduction in labor supply, especially among female headed households. They estimate a fairly imprecise and large reduction in labor hours (-658 hours per year, with a 95% CI of [-1186 , 130]; see Column (2) of Table 2 on page 157). They also estimate a large and imprecise change in annual earnings of -$2,943 (95% CI of [-10,169 , 4,284]). Corresponding to the tax rates operating around 1970, I assume a linear marginal tax rate of 20% on earnings, consistent with the absence of an EITC program during this time period. I arrive at 20% using the 14% bottom tax bracket for federal taxes and a 6% state tax assumption. With this assumption, the net change tax revenue collected due to behavioral responses to food stamps is $2,943*.2=$588.60. It is important to note that this estimate is not statistically significantly different from zero.

In contrast, the food stamp program provided an average monthly benefit of $26.77 per per-

policy elasticity would be $FE^{EITC} = 0.09$, ranging between 0.07 and 0.11.

[57] This is potentially an overestimate of the net effect of behavioral responses because some of the responses found in Chetty et al. (2013) is along the extensive margin and is more amenable to the potential critique that the earlier literature could not effectively separate the impact of EITC expansions from the impact of the decrease in welfare generosity (see Meyer and Rosenbaum (2001) for this debate). Therefore, I also consider the case that the 0.05 figure in Chetty et al. (2013) captures all of the EITC response (so that $FE^{EITC} = 0.05$). This arguably provides a lower bound of the impact of the policy. For an upper bound, I consider the upper range of extensive margin response can be added to Chetty et al. (2013), so that $FE^{EITC} = 0.11 + 0.05 = 16\%$.

son in 1978[58], which corresponds to \$321.24 per person per year. Hoynes and Schanzenbach (2009) estimate a mean household size of 3.59 in their sample, which implies a household-level transfer size of \$1,153.25. Hence, the total cost to the government of providing the food stamps policy is \$1,153.25+\$588.60=\$1,741.85.

For the net valuation of food stamps, Smeeding (1982) estimates that food stamps are valued dollar-for-dollar. In contrast, Whitmore (2002) estimates that every dollar of food stamps is valued at ~\$0.80 by the beneficiaries. In the absence of behavioral responses this estimate suggests the MVPF would be 0.8. Placing this into the context of the size of the transfers, the estimate suggests that the mechanical transfer of \$1,153.25 is valued by beneficiaries at only \$922.60.

F.3 Section 8 Housing Vouchers

Jacob and Ludwig (2012) study the impact of obtaining a housing voucher on labor supply (intensive + extensive), Medicaid receipt, TANF receipt, and SNAP receipt. For the extensive margin labor supply response, I use the 11% tax rate assumption from the EITC section. For the intensive margin response, Jacob and Ludwig (2012) report a marginal tax rate of 24% for the treatment group that includes phase-out of government benefits in addition to marginal income tax rates. For the change in TANF and SNAP use, I use the Green Book (2004) and compute average costs per household in 2002, normalized to 2007 dollars using the CPI-U to be consistent with Jacob and Ludwig (2012). For the change in Medicare enrollment, I use costs compiled by Holahan and McMorrow (2012). Table A1 reports the calculations.

[58] www.fns.usda.gov/pd/SNAPsummary.htm+&cd=3&hl=en&ct=clnk&gl=us&client=safari

Table A1: Housing Voucher Calculation

Item	Value	Source
Voucher Cost	$8,400	Jacob and Ludwig (2012)
Extensive Margin		
Extensive Margin Change	0.036	
Mean earnings	22,232	Mean earnings of $5557 per quarter (Jacob and Ludwig (2012))
Assumed Avg Tax Rate	-11%	Avg EITC subsidy (see EITC section)
Fiscal Impact	-88.0	
Intensive Margin		
% Intensive	0.592	Jacob and Ludwig (2012)
Earnings change	910	Jacob and Ludwig (2012) report quarterly change of 227.54
Assumed Tax Rate	24%	Jacob and Ludwig estimate 18% tax rate in control group and 24% tax rate in treatment group that includes phase-out of services
Fiscal Impact	129.3	
TANF		
% Increase	0.017	Jacob and Ludwig (2012)
Avg Cost	254	Mean monthly TANF from Green Book (2007 dollars deflated using CPI-U as in Jacob and Ludwig (2012))
Fiscal Impact	4.3	
Medicaid		
% Increase	0.058	Jacob and Ludwig (2012)
Avg Cost per enrollee	6192	Halahan and McMorrow (2012) Appendix Table 1 reports costs for 2002
Fiscal Impact	359.1	
Food Stamps		
% Increase	0.076	
Avg Cost per family	364	Family of 4 assumption
Fiscal Impact	27.7	
Total Program Cost	$8,832.4	
Total Behavioral Impact	$432.4	
% of cost	0.05	
Valuation of $1 of Voucher	0.83	Reeder (1985)
Value of $8,400 Voucher	6,972	
Net Valuation	-1,428	
% of cost	-0.16	
MVPF	$0.79	