Online Appendix for "Can Fiscal Externalities Be Internalized"

Erzo F.P. Luttmer

June 27, 2022

1 Overview of the online appendix

In this online appendix, I underpin the verbal arguments made in the main paper with formal models. Section 2 lays out a reference case in which all sponsors are passive, in which case welfare is the same as in a world without sponsors. The next three sections show how sponsors achieve welfare gains over the references case in three different settings. Section 3 describes how sponsors can offer individuals incentives for human capital investments, and the resulting effects on social welfare. Section 4 adopts the optimal nonlinear income-tax framework from Mirrlees' (1971)[3], and shows under what conditions sponsors can achieve welfare gains in that setting. Finally, Section 5 demonstrates how the previous two models, after reinterpretation of some variables, also apply to social-insurance settings.

2 Status quo with passive sponsors

If sponsors do not innovate or take any action, they do not affect individuals. The outcome for individuals would in that case be identical to the outcome for individuals in the absence of sponsors. Because individuals are not affected, their fiscal externalities are also not affected. Under the assumption that sponsors are risk neutral and administrative costs are negligible, sponsors are indifferent between whether or not to sponsor randomly matched individuals when the group-specific adjustment term is zero (i.e, $\phi_g = 0$ in the equation for the sponsor fee in the main paper) because, in that case, the payment the sponsor receives from the government for each sponsored individual (i.e., the fiscal externality from this individual) is equal in expectation to the sponsor fee it must pay (the fiscal externality of the average individual sponsored by other sponsors). Hence, if all sponsors are passive, the expected surplus of sponsors that are matched to individuals is zero

and the expected effect on the government budget is zero. Hence, social welfare when all sponsors are passive is equal to social welfare in the absence of sponsors.

In the online appendix, I drop the subscript t from fiscal externalities and from the transfer fee because all the formal models presented here are static. Moreover, I drop the subscript g for groups on which matching was conditioned because such conditioning is not used in the formal models. Using superscript 0 to denote outcomes in the status quo with passive sponsors and the expectations operator to denote expectations across different individuals to which the sponsors may be matched, we have the following:

Lemma 1. *(Status quo with passive sponsors). If sponsors are risk neutral and administrative costs are negligible, the status-quo outcome in which all sponsors are passive is characterized by the following: (i) an adjustment term in the formula for the sponsor fee of $\phi^0 = 0$; (ii) an expected sponsor fee equal to the expected fiscal externality, $\bar{T^0} = E[F_{ij}^0]$; (iii) all individuals being equally well off as in a setting without sponsors; (iv) all sponsors having an expected surplus of 0; and (v) social welfare being the same as in a setting without sponsors.*

In short, having sponsors has no effects on individuals or social welfare if all sponsors are passive and risk neutral and if administrative costs are negligible. In the next sections, we can therefore compare outcomes relative to a status quo with passive sponsors, because lemma 1 establishes that the outcome of this status quo is the same for individuals and in terms of social welfare as the outcome that occurs in the absence of any sponsors. The next three sections present three models that show how actions by sponsors increase welfare over the status quo in three different settings.

3 Human-capital investments

Setup.

Individuals have different characteristics that affect their returns to human-capital investments. Let the set of a person's characteristics be denoted by $\theta \in \mathbb{R}^n$. Examples of elements of θ might be characteristics such as risk aversion, writing ability, physical strength, dexterity, patience, emotional intelligence, knowledge of geometry, or carpentry skills. In short, the elements of θ include all individual-specific determinants of the return to a particular human-capital investment, be it a college course in philosophy or participating in an Outward Bound program. Therefore, thinking of n as being very large, and the elements of θ perhaps being even more detailed than the examples mentioned above, makes sense.

Similarly, many types of human-capital investment are possible, for example, different academic courses, internships, technical training, coaching, leadership or social-skills training, and so

forth. Moreover, they would be distinguished by the way they are taught or run. Thus, the vector of possible human-capital investments, $I \in \mathbb{R}^m$, is also highly dimensional. Not all types of potential investments are currently known. The number of known types of investments is $\tilde{m} < m$, and people can invest only in known types of investments. The number of known types expands whenever new types of investments are invented or discovered.

The private and social return to any combination of these investments may depend on the individual's current level of human capital and other characteristics, as described by θ. The return of an investment that is captured by the individual minus the cost to the individual of the investment is denoted by the function $R(\theta, I) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. These net returns may reflect monetary returns such as higher earnings but also intangible returns such as more pleasant work conditions, better personal relationships, or less involvement with the criminal justice system. Similarly, the net return is lowered both by the monetary costs of the investment, such as tuition or foregone earnings, and by intangible costs, such as stress, effort, or foregone leisure. The net return may be uncertain due to various individual-specific and economy-wide shocks. I do not explicitly model this uncertainty, but instead take it into account by defining R as the person's ex-ante equivalent variation of the investment. Hence, the net return is a risk-adjusted net return.

A person's fiscal externality is their net effect on the government budget. Thus, payments of taxes or fees by a person to the government increase that person's fiscal externality, and receipts of government transfers reduce it. Similarly, the government's spending on a person, such as educational spending or spending through the criminal justice system, make a person's fiscal externality less positive or more negative. The expected effect of investment I by a person with characteristics θ on the person's fiscal externality is denoted by the function $\Delta F(\theta, I) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

Outcome without sponsors.

In a setting without sponsors, individuals may fail to invest optimally in their human capital for three main reasons.

First, finding the combination of human-capital investments I that maximizes the net return R for an individual with characteristics θ is a highly multidimensional problem, which is hard to solve. Whether an individual (or their parents), who only solves this problem at most a few times in their lifetime, is able to find the optimum or something close to the optimum is unclear. Of course, individuals may learn from others, but other people have different characteristics, and outcomes are only observed with a lag. They may also obtain advice from various sources (counselors, deans, college guides, etc.), but whether these sources have strong incentives to provide personalized and accurate advice is unclear. Although many people clearly spend a lot of time and energy on making human-capital decisions, these decisions are hard to get right.

Second, individuals don't take into account the fiscal externalities of their human-capital deci-

sions. Although fiscal externalities for some individuals may be small on net, for example, if educational costs paid by the state are approximately offset by subsequent higher tax revenues from the individual's higher earnings, fiscal externalities are generally not zero and they vary across individuals. Thus, individuals maximize R rather than $\Delta F + R$, which is the effect on social welfare. Fiscal externalities can potentially be very large, for example, if poor school quality induces teenagers to drop out and become involved with crime.

Third, individuals may face liquidity constraints. The financial sector and the government have various lending and grant programs to reduce liquidity constraints. Still, sponsors have an advantage over the financial sector in reducing liquidity constraints that individuals face because they receive the fiscal externalities of educational investments, and therefore have stronger incentives to ensure educational investments with positive net social value are not forgone.

Incentives for a sponsor to contract.

Because the government pays the sponsor the fiscal externalities of the sponsored individual, the sponsor internalizes all fiscal externalities from the individual's human-capital investments. The individual receives the net private returns from their human-capital investments. Hence, the sponsor and the sponsored individual together capture the entire welfare gain associated with these human-capital investments. If the conditions of the Coase Theorem apply, contracting between the sponsor and sponsored individual results in the socially efficient level of investment. Such contracting should be feasible and relatively low cost because it involves just two contemporaneous parties. However, information asymmetries, such as the individual having private information about net private returns, could prevent the individual and their sponsor from fully reaching the efficient outcome. Still, any contract to which the sponsor and the sponsored individual agree allows both parties to do better than the status quo, and improves social welfare. Sponsors that are better able to overcome any information problems earn higher surplus and outcompete other sponsors in being matched with individuals.

To establish this result more formally, let the vector I_{ij} denote the human-capital investments that individual i undertakes when matched to sponsor j but would not have undertaken in the absence of a sponsor. The effect of these investments on social welfare, ΔW_{ij}, is equal to their effect on the individual's private net return, R_{ij} $(\equiv R(\theta_i, I_{ij}))$, plus their expected effect on the individual's fiscal externality, ΔF_{ij} $(\equiv \Delta F(\theta_i, I_{ij}))$.

Sponsor j induces individual i to undertake the incremental human-capital investments by offering an incentive schedule $S_{ij}(I) : \mathbb{R}^m \to \mathbb{R}_0^+$. The realized incentive payment, given the individual's choice of investment, is denoted by $S_{ij} \equiv S(I_{ij}) \geq 0$. The total effect of these human-capital investments on individual i's surplus is therefore $R_{ij} + S_{ij}$, and a rational individual agrees to undertake these human-capital investments only if $R_{ij} + S_{ij} \geq 0$. The sponsor fee that sponsor j

pays to the government, $\bar{T}_j$, is not affected by the incremental human capital investments of person i, because the sponsor fee depends on the fiscal externalities of individuals not sponsored by sponsor j. Given that the government pays sponsor j the fiscal externalities of individual i, the total effect of the human-capital investments on sponsor j's expected surplus is $\Delta F_{ij} - S_{ij}$. A risk-neutral and surplus-maximizing sponsor offers incentive payment S_{ij} for incremental investments I_{ij} only if $\Delta F_{ij} - S_{ij} > 0$. The effect of the human-capital investments on the joint surplus of the sponsor and the sponsored individuals is therefore equal to their effect on social welfare, that is, $(R_{ij} + S_{ij}) + (\Delta F_{ij} - S_{ij}) = \Delta W_{ij}$. Moreover, because $R_{ij} + S_{ij} \geq 0$ and $\Delta F_{ij} - S_{ij} > 0$ for any contract to which the individual and sponsor both agree, the resulting effect on social welfare must be positive.

Outcomes when a sponsor innovates.

In addition to increasing surplus through Coasian contracting, a sponsor may also discover new types of human-capital investment that increase welfare. At an expected cost of C_{ij}, sponsor j can expand the set of human-capital investments available to individual i. For example, the new type of investment could be a new type of on-the-job training program or a better way of assessing an individual's fit for different types of jobs. Formally, we can think of the discovery as increasing the space from which investments I are selected, and thus as an increase in $\tilde{m}$ that is specific to individual i and sponsor j. The cost and welfare gain of discovering new types of human-capital investment may vary across sponsors because different sponsors have different skills in discovering new ways of investing in human capital.

As in the previous subsection, the sponsor and the sponsored individual jointly capture the entire effect on social welfare of their actions, which now consist not only of the individual's choice of investments I_{ij} given the incentive payment S_{ij} that the sponsor and individual agreed upon, but also the sponsor's choice of spending C_{ij} to expand the set of feasible human-capital investments. Specifically, social welfare increases by $\Delta W_{ij} = R_{ij} + \Delta F_{ij} - C_{ij}$ as a result of the innovation by the sponsor and the investment in human capital by the individual. The individual's surplus weakly increases by $R_{ij} + S_{ij} \geq 0$ and the sponsor's expected surplus increases by $\Delta F_{ij} - S_{ij} - C_{ij} > 0$.

The setup with sponsors therefore allows a sponsor and their sponsored individuals to jointly capture the entire effect on social welfare of innovations that create new types of human-capital investments. Currently, private parties don't capture the entire effect of such innovations, because they do not internalize the fiscal externalities of investments made possible by these innovations. The setup with sponsors therefore aligns incentives for innovation more closely with social welfare. The socially optimal level of innovation may nevertheless not be achieved for the same reasons that also apply in product markets without externalities. The sponsor may not capture the entire increase in surplus, because it ends up sharing some of the rents with sponsored individuals (if

they are heterogeneous or have private information), just like in product markets an innovator may not be able to extract the entire increase in consumer surplus when consumers are heterogeneous. If the innovation is non-rival and non-excludable, the innovator may also not receive the surplus from other sponsors adopting the innovation. These two forces result in incentives for a sponsor to innovate that are below the socially optimal level. On the other hand, if the sponsor's innovation leads to a transfer of rents from other sponsors to this sponsor, the sponsor may have incentives for innovation that are greater than the socially optimal level. This result is analogous to the "business-stealing" effect on innovation in traditional product markets.

Proposition 1. (Effects of Innovation and Contracting by a Single Sponsor) *If (i) the sponsored individual is rational, (ii) the sponsor maximizes its expected surplus, (iii) the sponsor is risk neutral, (iv) the cost of measuring fiscal externalities and administering sponsor fees is negligible, (v) no non-fiscal externalities exist, and (vi) the sponsor can provide only nonnegative incentives, then holding constant the behavior of other sponsors and individuals, any outcome resulting from innovation and contracting between the sponsor and an individual it sponsors (i) weakly increases the risk-adjusted surplus of the individual, (ii) increases the expected surplus of the sponsor, (iii) increases social welfare, and (iv) increases the expected fiscal externality (defined as positive for positive externalities) exerted by the sponsored individual.*

Proof. Consider the incentives for and effects of an innovation of sponsor j sponsoring individual i, under the Nash assumption that the behavior of all others is held constant. The sponsor spends C_{ij} to expand the set of feasible human-capital investments for individual i, and offers incentive payment S_{ij} to the individual i to undertake human capital investments I_{ij} that the individual would not have taken otherwise. A rational individual agrees to take up the sponsor's incentive payment only if their surplus weakly increases, $R_{ij} + S_{ij} \geq 0$. Next, a sponsor that maximizes expected surplus spends C_{ij} and offers incentive payment S_{ij} only if doing so increases their expected surplus. Hence, innovation and contracting increases the sponsor's expected surplus, $\Delta F_{ij} - S_{ij} - C_{ij} > 0$. Because no non-fiscal externalities exist, innovation and contracting do not affect the surplus of any of the other sponsors or other individuals. Given that the payment of the government to sponsor j is equal to the fiscal externality of individual i and given that the sponsor fee of sponsor j depends only on the fiscal externalities of individuals sponsored by other sponsors (and those fiscal externalities are not affected, because the behavior of others is held constant and because no non-fiscal externalities exist), no net effect of the innovation or investment on the government's budget arises.[1] Hence, the effect on social welfare is the sum of the effects on the surplus of individual i and the expected surplus of sponsor j, $\Delta W_{ij} = (R_{ij} + S_{ij}) + (\Delta F_{ij} - S_{ij} - C_{ij}) > 0$. The

[1] The increase in the expected fiscal externality of individuals sponsored by sponsor j increases the sponsor fee that other sponsors have to pay to the government. However, because this fee is a transfer payment between risk-neutral parties, it has no effect on social welfare.

welfare effect is positive because the effect on the surplus of the individual is nonnegative and the effect on the expected surplus of the sponsor is positive. Finally, given that the innovation cost is nonnegative ($C_{ij} \geq 0$) and given that the sponsor can only make nonnegative incentive payments ($S_{ij} \geq 0$), the result that expected sponsor surplus increases implies that the expected effect on the fiscal externality is positive: $\Delta F_{ij} - S_{ij} - C_{ij} > 0 \Rightarrow \Delta F_{ij} > S_{ij} + C_{ij} \geq 0$. □

Equilibrium outcome when sponsors compete and innovate.

The previous subsection considered the effects of innovation and contracting by a single sponsor. This subsection analyzes the effects when multiple sponsors compete to sponsor a given set of individuals and multiple sponsors engage in innovation and contracting.

By Proposition 1, social welfare increases as a result of contracting and innovation of sponsor j and sponsored individual i if the behavior of other sponsors (and therefore the sponsor fee T_j) is held constant. However, changes in sponsor fees do not affect total welfare, because they are transfer payments between sponsors and the government. Thus, by viewing the contracting and innovations of different sponsors for different sponsored individuals as occurring sequentially (over the set of sponsors that end up being a sponsor in equilibrium), with the behavior of others being held constant when the contracting and innovation for a particular sponsor and sponsored individual occurs, it follows that social welfare also increases when multiple sponsors innovate.

Given that the matching of individuals to sponsors is random, the entry decision of a sponsor is determined by the expected surplus it expects to achieve for individuals who are randomly chosen (rather than for the individuals with whom it was matched ex post). If entry is free and if each sponsor is sufficiently small that it does not consider the effects of its entry on the sponsor fee, expected sponsor fees must adjust in equilibrium until the marginal sponsor, denoted by $\hat{j}$, has an expected surplus of zero. The proof of Proposition 1 establishes that if the sponsor fee is held constant, contracting and innovation increase the expected surplus that sponsor j receives from being matched to individual i by $\Delta F_{ij} - S_{ij} - C_{ij} > 0$. This result implies that if the sponsor fee is held constant, the expected surplus of the marginal sponsor, denoted by subscript $\hat{j}$, increases by $E[\Delta F_{i\hat{j}} - S_{i\hat{j}} - C_{i\hat{j}}] > 0$, where $E[.]$ denotes expectations over individuals i that could be randomly matched to this sponsor. Because competition between sponsors drives the expected surplus for the marginal sponsor to zero, the sponsor fee must in equilibrium increase in expectation by $E[\Delta F_{i\hat{j}} - S_{i\hat{j}} - C_{i\hat{j}}] - \bar{T^0}$, where $\bar{T^0}$ denotes the expected sponsor fee associated with the passive outcome. Hence, sponsors get to keep only the inframarginal expected surplus from contracting and innovation, with the remainder of their surplus being transferred to the government via higher sponsor fees. These results are summarized in the following proposition:

Proposition 2. (Equilibrium Outcome with Sponsor Innovation and Contracting) *If (i) the assump-*

tions of Proposition 1 hold, (ii) there is free entry by sponsors, and (iii) sponsors are sufficiently small that competition between them drives the expected surplus of the marginal entrant to zero, innovation and contracting between the sponsors and their sponsored individuals results in an equilibrium increase with higher social welfare than in the equilibrium without sponsors. The welfare gains from contracting and innovations are shared between individuals, the government (through higher sponsor fees), and inframarginal sponsors.

Without imposing further structure, determining how the welfare gains are distributed is impossible beyond establishing that the marginal sponsor captures no surplus and that the government's revenue increases. More of the welfare gains go to individuals if individuals have more bargaining power or private information that results in information rents. If human-capital investments affect skill-specific wages, changes in the wage structure affect the distribution of surplus across individuals. Inframarginal sponsors capture a greater share of the welfare gains if dispersion among sponsors in the returns from innovation and contracting is greater. In the extreme case in which sponsors are identical, they don't capture any of the welfare gains. Whatever welfare gains are not captured by individuals or inframarginal sponsors go initially to the government in the form of higher sponsor fees. The government can then distribute this surplus as it sees fit, for example, by lowering tax rates or by increasing transfers to lower-income people.

4 Income taxation – better screening

The model of the role of sponsors in human-capital investment focused on sponsors discovering new types of investments that yield a higher social return than current investments. The model also recognized the role of sponsors in improving human-capital investment decisions that might be suboptimal for other reasons, including a wedge between private and social returns, liquidity constraints by individuals, and lack of information on the returns to investments by individuals. The model of the role of sponsors in income taxation emphasizes information asymmetries between individuals and the government, and describes how sponsors may be able to reduce the resulting inefficiencies.

Setup.

The setup follows the standard Mirrlees' (1971)[3] optimal income-tax model, except where noted otherwise. I normalize the population to 1. Individuals have utility functions defined over consumption c and labor supply l: $u_i(c,l)$ with $\partial u_i/\partial c > 0$, $\partial^2 u_i/\partial c^2 < 0$, $\partial u_i/\partial l < 0$, and $\partial^2 u_i/\partial l^2 < 0$. The utility function is indexed for each individual i because, unlike Mirrlees (1971)[3], I don't assume the utility function is identical for all individuals. Each individual's abil-

ity level $w \geq 0$ is fixed and is private information, but the utility function is common information. Abilities are distributed according to the density function $g(w)$. Individual earnings $z \equiv wl$ are observable and determine the individual's tax liability through the tax function. Given that the tax an individual pays is a fiscal externality, the tax function is denoted by $F : \mathbb{R}_0^+ \to \mathbb{R}$ for consistency with the notation used in the previous section. Thus, an individual with earnings z pays $F(z)$ in taxes to the government, and the individual's budget constraint is given by $c = z - F(z)$. Individuals choose c and l to maximize their utility subject to the budget constraint. The resulting choices depend on the individual's privately observed ability level w and are denoted by: $c_i(w)$ and $l_i(w)$. The individual's utility is $v_i(w) \equiv u_i(c_i(w), l_i(w))$. For the moment, I do not define the government's objective or require that the tax function is set to maximize the government's objective function. Instead, the only restrictions on the tax function are that $0 \leq F'(z) \leq 1$. Payments between the government and sponsors and the way sponsors are matched to individuals are as described in the main paper. In particular, the government pays the sponsor of an individual the tax that the individual pays to the government, $F(z)$, because the tax is the individual's entire fiscal externality in this setup. In return, the sponsor pays the government a sponsor fee as defined in the main text.

Case 1: A perfectly informed sponsor.

Sponsors could better observe, or better be able to use, information about individuals' abilities than the government in many ways. For example, people in practice may have a pretty good idea of other individuals' abilities, but this information may not be contractible. Sponsors can use such non-contractible information, whereas governments have not been able to use such information for taxation. Sponsors may also be able to use different tools or devise new ways to get a more accurate estimate of someone's ability. For example, a sponsor that employs an individual may use different types of work assignments to better distinguish individuals with different abilities. Or sponsors may develop and measure various proxies for ability. In short, the private sector in many ways likely has advantages over the government in ascertaining any given individual's ability.

To illustrate how sponsors could reduce inefficiencies related to asymmetric information, first considering the extreme but unrealistic case in which sponsors can perfectly observe individuals' abilities is useful. The next subsection considers the more realistic case in which sponsors have an information advantage over the government but still cannot perfectly observe individuals' abilities. The benefit of first considering the extreme case of sponsors' perfectly observing ability is that it clearly shows how the government determines the degree of redistribution, what the sources of the efficiency gains are, and what incentives sponsors have to realize these efficiency gains.

Individuals have the outside option of not engaging with their sponsor and receiving the same level of utility as they would receive in the absence of sponsors. This level of utility is denoted by $v_i^0(w)$, and the government's tax schedule determines the distribution of $v_i^0(w)$ across individuals.

However, the tax schedule creates efficiency losses because individuals choose labor supply subject to a marginal tax $F'(z)$, but $F'(z) > 0$, at least for some individuals. Because a sponsor receives the tax of the individual it sponsors, the sponsor and the individual jointly capture the entire change in surplus from a change in labor supply. The sponsor can induce the individual to supply the efficient amount of labor by offering the person an individual-specific incentive schedule equal to $S_i(z) = \max(F(z) + \alpha_i(w), 0)$, where $\alpha_i(w)$ solves $\text{argmax}_l \left(u_i\left(wl + \alpha_i(w), l\right)\right) = v_i^0(w) \quad \forall w$.

In the relevant range (i.e., when $F(z) + \alpha_i(w) \geq 0$), the incentive schedule pays the individual $F(z)$, which is the tax that the individual pays to the government, plus an an individual-specific constant term $\alpha_i(w)$. On the margin, this incentive schedule therefore fully offsets the income tax paid by the individual, thereby providing the individual with undistorted incentives for labor supply. A rational individual accepts this incentive schedule because the individual-specific constant term $\alpha_i(w)$ is defined such that the incentive schedule provides the individual with the same level of utility as before, $v_i^0(w)$. Hence, the distribution of utilities across individuals is still determined by the tax schedule and is not affected by the incentive schedules offered by sponsors. In other words, the government remains in control of the degree of income redistribution.

Under the incentive schedule, the individual's budget constraint simplifies to $c = z - F(z) + S_i(z) = z + \alpha_i(w)$ in the relevant range of earnings. The term z in the budget constraint shows that the individual keeps all their labor earnings on the margin, and thus faces undistorted labor-supply incentives. However, the sponsor captures inframarginal tax revenues through the individual-specific constant term $\alpha_i(w)$, which acts as an individual-specific lump-sum tax from the perspective of the individual. The constant term $\alpha_i(w)$ is also equal to minus the equivalent variation of the tax system (by definition of an equivalent variation). Whenever the income tax distorts labor supply (i.e., when $F' \neq 0$), it creates a deadweight loss equal to the equivalent variation minus the tax revenue: $DWL_i = -\alpha_i(w) - F(z_i^0) > 0$, where z_i^0 is the level of earnings that individual i would have chosen under the tax schedule $F(z)$ if they were not matched to a sponsor. The incentive schedule eliminates this deadweight loss. Figure (1) depicts the budget constraints, with sponsor incentives (solid red) and without sponsor incentives (blue), for two individuals. The figure also shows the level of earnings that the individuals (denoted by subscripts 1 and 2) choose without sponsor incentives (z_1^0, z_2^0) and with sponsor incentives (z_1^1, z_2^1). Finally, the figure shows the individuals' indifference curves (green) and the DWL that is eliminated when sponsors provide each individual with the optimal individual-specific incentive schedule.

If sponsor j were passive, it would receive from the government $F(z^0)$, the individual's fiscal externality under the tax schedule, and pay the sponsor fee $\bar{T}_j$ to the government. Thus, if passive, the sponsor's surplus would be $F(z^0) - \bar{T}_j$. When sponsor j offers incentive schedule $S_i(z)$ to the individual, its surplus is $F(z^1) - \bar{T}_j - S_i(z^1) = \alpha_i(w) - \bar{T}_j$. Hence, at a given sponsor-fee level, offering the incentive schedule increases the sponsor's surplus by $\alpha_i(w) - F(z^0) > 0$, that is, by

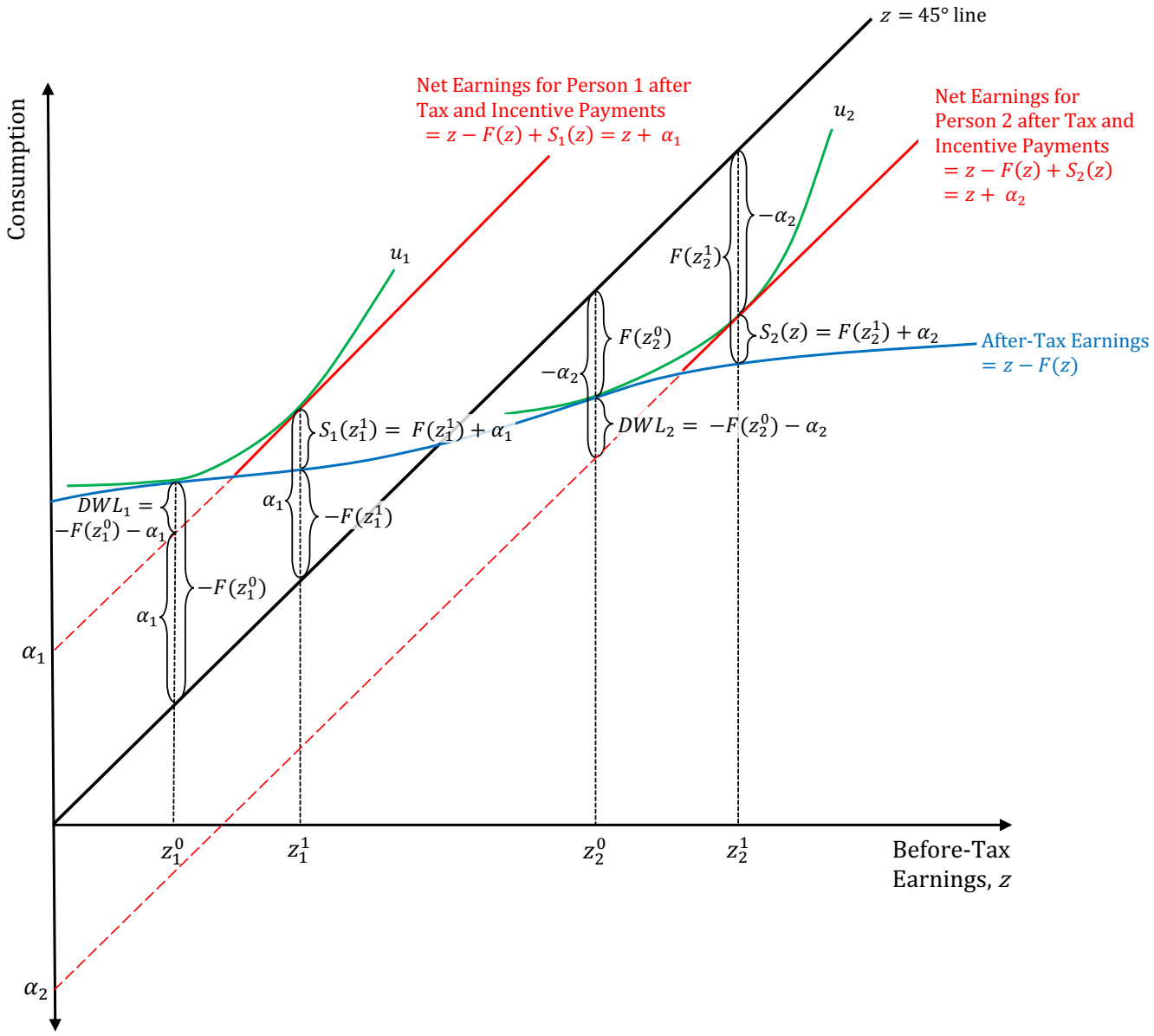

Figure 1: Outcome with Sponsors having Perfect Information

the amount of deadweight loss created under the tax schedule if the sponsor were passive. Hence, surplus-maximizing sponsors offer these incentive schedules. As explained in the previous section, the government sets the adjustment term ϕ in the formula for the sponsor fee such that the marginal sponsor earns an expected surplus of zero from being matched to a random individual. Hence, the sponsor fee will increase by the deadweight loss eliminated by the marginal sponsor's incentive schedule. The government can use the increase in sponsor-fee receipts to lower the tax schedule and thereby return to individuals some of the surplus generated from eliminating deadweight loss in labor supply. Sponsors that eliminate more deadweight loss than the marginal sponsor retain this excess of deadweight loss eliminated as surplus.

Case 2: A sponsor receiving an imperfect signal of individual ability.

In the case described above, the strong assumption that a sponsor can perfectly observe individual abilities w made it possible to design an incentive schedule that eliminates all labor-supply distortions. Although a sponsor has many ways in which it can obtain better information about individuals' abilities than is available to the government, its information about individuals' abilities will in practice not be perfect. This subsection relaxes the assumption that a sponsor can perfectly observe individuals' abilities. The better information obtained by a sponsor is instead modeled as the sponsor receiving an informative signal of an individual's ability. I prove below that if this signal is sufficiently strong, the sponsor can design an individual-specific and positive incentive schedule that reduces inefficiency, makes the individual weakly better off, and increases the sponsor's surplus. The intuition behind the efficiency improvement is that the schedule partially offsets existing tax disincentives for labor supply at those earnings levels that the signal indicates are most likely for the individual in question. In short, the incentive schedule creates work incentives that are better targeted to an individual's likely earnings levels. The incentive schedule makes the individual weakly better off because a sponsor is allowed to give individuals only nonnegative incentives; that is, an incentive schedule can give an individual extra money but cannot take money away from them. Finally, the surplus of the sponsor increases if the signal is sufficiently strong. In that case, the sponsor can target the schedule sufficiently well that the increase in labor supply due to the stronger incentives creates a sufficiently large positive fiscal externality, which accrues to the sponsor and offsets the cost to the sponsor of making the incentive payment. As before, competition between sponsors would in the longer run result in an increase in the sponsor fee, thereby reducing the surplus of sponsors. The government can use the increase in sponsor-fee revenue to further increase individuals' utility.

An incentive schedule as a perturbation to the tax function.

The proof that a sufficiently informative signal of individual ability increases the surplus to the sponsor is based on the perturbation techniques introduced by Saez (2001)[4]. I maintain his assumptions and largely follow his notation.[2] $H(z)$ denotes the cumulative distribution function of earnings z, with associated density function $h(z)$, at the current tax schedule. The sponsor receives a signal of each individual's ability. Let the density function of earnings conditional on signal k be denoted by $\hat{h}_k(z)$. I assume the signal is informative, that is, that $h(z) \neq \hat{h}_k(z)$ for some value of z. Following Saez (2001)[4], $h^*(z)$ denotes the *virtual* density, which is the density of earnings that would have occurred if the tax function $F(z)$ were linearized around point z.[3] Similarly, the conditional virtual density function is denoted by $\hat{h}^*_k(z)$. Because perturbations to a non-optimal tax schedule could increase the surplus of the sponsor even if the sponsor receives no signal at all, I conservatively assume the tax function $F(z)$ is set optimally given society's preferences for redistribution. These preferences are described by the function $g(z)$, which gives, at the optimum, the marginal social value of consumption for individuals with earnings z relative to the marginal value of public funds.

I hold the tax function fixed and consider a perturbation in labor supply stemming from the sponsor offering an incentive schedule $S_k(z)$ to an individual sponsored by the sponsor and about whom the sponsor receives signal k. The sponsor offers this individual an incentive schedule that pays zero until earnings level $z^*_k - dz^*_k$, increases with slope $ds_k > 0$ for earnings between $z^*_k - dz^*_k$ and dz^*_k, and then remains constant at $ds_k\, dz^*_k$ for earnings above z^*_k. I consider a change in slope ds_k that is second order relative to dz^*_k, the length of the interval over which it applies. Figure 2 illustrates the tax function and this incentive schedule. The incentive schedule is the mirror image of the perturbation to the tax schedule considered in Saez (2001)[4]; that is, rather than increasing the marginal tax over the interval dz^*_k, marginal work incentives are increased over this interval. Consumption, or net earnings, for those with signal k is equal to pre-tax earnings z minus tax payment $F(z)$ plus incentive pay from the sponsor $S_k(z)$, as indicated by the dashed red line in the figure. Thus, consumption equals $c = z - F(z) + S_k(z)$.

The effect of offering this incentive schedule conditional on signal k on the sponsor's expected surplus consists of a mechanical effect, X^M_k, and the fiscal externality from the individual's labor-supply response to the incentive schedule.[4] The part of the fiscal externality that stems from the

[2]Specifically, I assume (i) leisure is a normal good, (ii) only local incentive constraints bind at the optimal tax schedule, and (iii) the curvature of each individual's utility function at their optimal earnings level is greater than the curvature of the tax schedule at that earnings level.

[3]Using the virtual density simplifies the notation and exposition because it allows us to ignore the effect of changes in the slope of the tax schedule that occur for small changes in labor supply.

[4]The expectation is taken over all possible realizations of the individual's earnings given signal k. The sponsor's realized surplus depends on the realized earnings of the individual it sponsors, because the incentive payment to the individual and the fiscal externality payment received from the government depends on realized earnings.

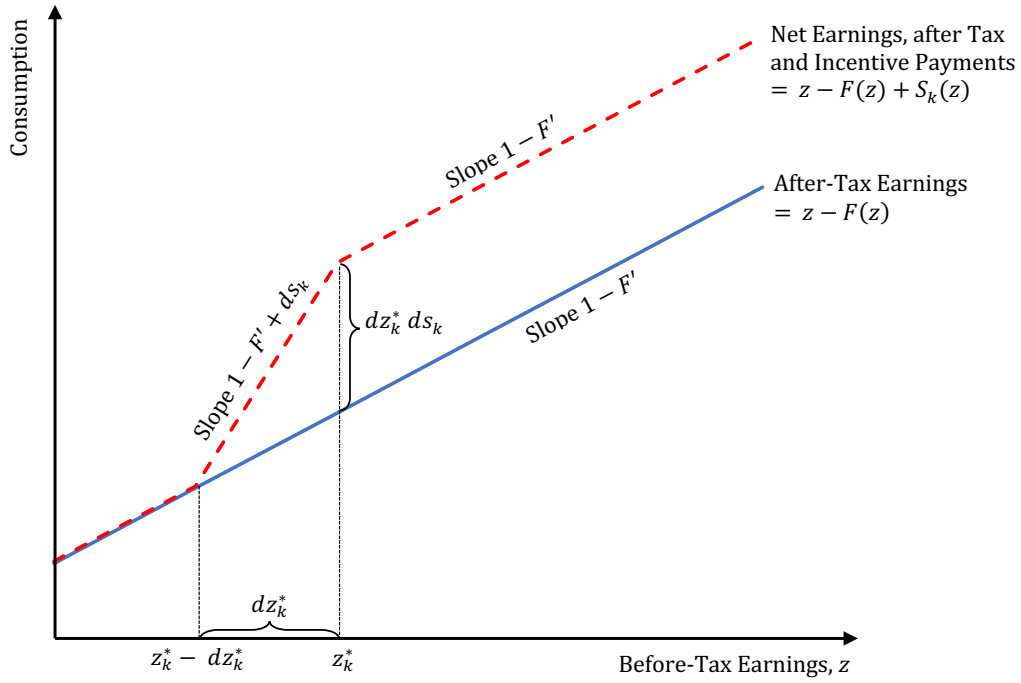

Figure 2: Perturbation due to the Sponsor's Incentive Schedule

substitution effect is denoted by X_k^S, and the part that stems from the income effect is denoted by X_k^I.

The mechanical effect reflects the expected cost to the sponsor of the incentive payments it would make if the individual kept their earnings constant.[5] The incentive schedule provides an income transfer of $ds_k \, dz_k^*$ for all earnings realizations above z_k^*.[6] The mechanical effect on the sponsor's expected surplus is therefore

$$X_k^M = -dz_k^* ds_k \int_{z_k^*}^{\infty} \hat{h}_k(z) dz.$$

Between earnings levels $z_k^* - dz_k^*$ and dz_k^*, the incentive schedule increases the individual's consumption by ds_k for each dollar of earnings. This labor-supply incentive increases earnings, which creates a fiscal externality of $F'(z)$ per dollar of increased earnings that accrues to the sponsor. The change in earnings due to the incentive schedule is determined by the compensated earnings elasticity, $\zeta^c_{(z^*)}$, at earnings level z_k^*, because the incentive schedule has no first-order effect on the level of utility, but it does have a first-order effect on marginal incentives.[7] The expected fiscal

[5] I define the mechanical effect as the effect purely on the sponsor's expected surplus. So, unlike the mechanical effect in Saez (2001)[4], it does not include the welfare effect of the transfer of income to the individual for earnings realizations above z_k^*, and it therefore does not depend on $g(z)$.

[6] The income transfer for earnings realizations between earnings levels $z_k^* - dz_k^*$ and dz_k^* is a third-order term, which can be omitted from X_k^M because the other term is a second-order term.

[7] The compensated earnings elasticity is defined as $\zeta^c \equiv \frac{1-\tau}{z} \frac{\partial z}{\partial (1-\tau)} |_{u=\bar{u}}$, where $\tau = F'(z)$ denotes the marginal tax

externality from the substitution effect is given by

$X_k^S = \left(\hat{h}_k^*(z_k^*)dz^*\right)\left(\frac{z_k^*}{1-F'(z_k^*)}\zeta^c_{(z_k^*)}ds_k\right)\left(F'(z_k^*)\right)$.

This expression consists of three terms, each between parentheses. The first term, $\hat{h}_k^*(z_k^*)dz_k^*$, is the probability of an earnings realization over the interval dz_k^*, where labor-supply incentives have increased. The second term, $\frac{z_k^*}{1-F'(z_k^*)}\zeta^c_{(z_k^*)}ds_k$, shows the increase in earnings, which is the compensated earnings elasticity, $\zeta^c_{(z_k^*)}$, scaled by $\frac{z_k^*}{1-F'}$ to turn it into the derivative of earnings with respect to the marginal incentive, times the change in the marginal earnings incentive, ds_k. The third term is the fiscal externality per dollar of increased earnings, which is equal to the marginal tax rate, $F'(z_k^*)$.

For an earnings realization above z_k^*, the individual receives a constant income transfer $dz_k^*ds_k$ from the incentive schedule.[8] This extra income changes labor supply through the income effect, and the change in labor supply affects the sponsor's expected surplus though its fiscal externality. The expected fiscal externality from the income effect is given by

$X_k^I = \int_{z_k^*}^{\infty}\left(dz_k^*ds_k\right)\left(\frac{\eta_{(z)}}{1-F'(z)}\right)\left(F'(z)\right)\hat{h}_k^*(z)dz$.

The expression again consists of three terms, each between parentheses. The first term, $dz_k^*ds_k$, is the increase in income due to the incentive schedule. The second term, $\frac{\eta_{(z)}}{1-F'(z)}$, shows the increase in earnings in reaction to a dollar of extra income from the incentive schedule. The definition of $\eta_{(z)}$, the parameter that defines the earnings response to income, follows the one used by Saez (2001)[4].[9] The third term is the fiscal externality per increase in earnings, which is equal to the marginal tax rate, $F'(z)$. These three terms are integrated over earnings levels for which the individual receives the incentive payment, namely, earnings above z_k^*, using the virtual distribution of earnings conditional on signal k.

The total change in the sponsor's expected surplus from the incentive schedule for an individual for whom it received signal k is given by

$X_k \equiv X_k^M + X_k^S + X_k^I = dz_k^*ds_k\left(-\int_{z_k^*}^{\infty}\hat{h}_k(z)dz + \frac{z_k^*F'(z_k^*)}{1-F'(z_k^*)}\zeta^c_{(z_k^*)}\hat{h}_k^*(z^*) + \int_{z_k^*}^{\infty}\frac{F'(z)}{1-F'(z)}\eta_{(z)}\hat{h}_k^*(z)dz\right)$.

Saez (2001)[4] shows the welfare effect of a perturbation to the tax schedule by $d\tau$ over earnings over an interval dz^* at earnings level z^* is given by[10]:

$X^{\text{welfare}} = dz^*d\tau\left(\int_{z^*}^{\infty}(1-g(z))h(z)dz - \frac{z^*F'(z^*)}{1-F'(z^*)}\zeta^c_{(z^*)}h^*(z^*) - \int_{z^*}^{\infty}\frac{F'(z)}{1-F'(z)}\eta_{(z)}h^*(z)dz\right)$.

The terms in X^{welfare} are the same as those in X_k, except that they are opposite in sign (because his perturbation is a tax increase rather than a decrease), the distributions are unconditional rather

rate and $u = \bar{u}$ indicates utility is held constant.

[8] The income transfer for earnings realizations between $z_k^* - dz_k^*$ and dz_k^* gives rise to a third-order term, which can be omitted from X_k^I because all other terms are second-order terms.

[9] $\eta \equiv (1-\tau)\frac{\partial z}{\partial I}$, where I denotes non-labor income and $\tau = F'(z)$ denotes the marginal tax rate. The Slutsky equation allows one to express η as the difference between the uncompensated and compensated earnings elasticities: $\eta = \zeta^u - \zeta^c$, where the uncompensated earnings elasticity is defined as $\zeta^u \equiv \frac{1-\tau}{z}\frac{\partial z}{\partial(1-\tau)}$.

[10] This expression is a rearrangement of equation (19) in Saez (2001)[4]. My notation is the same as his except that I use $F(.)$ to denote the tax function, whereas Saez uses $T(.)$.

than conditional, and the mechanical term includes $-g(z)$. Because X^{welfare} measures the welfare effect of a perturbation, it values the decrease in income transfers for earnings realizations above z^*, which is captured by $-g(z)$. By contrast, X_k does not measure a welfare effect but measures the effect of the incentive schedule on the sponsor's expected surplus. It therefore does not value how the incentive schedule transfers income to the individual for earnings realizations above z^*.

For an optimal tax schedule, the welfare effect of any perturbation needs to be zero, so $X^{\text{welfare}} = 0$. In particular, the welfare effect remains zero for a perturbation of $d\tau = ds_k$ over an interval of $dz^* = dz_k^*$ at earnings level $z^* = z_k^*$. Adding the expression for X^{welfare} for this particular perturbation to the expression for X_k, and using $\Delta h_k(z)$ to denote the difference between the conditional and unconditional density (so $\Delta h_k(z) \equiv \hat{h}_k(z) - h(z)$ and $\Delta h_k^*(z) \equiv \hat{h}_k^*(z) - h^*(z)$), yields

$$X_k = dz_k^* ds_k \Big(\tfrac{z_k^* F'(z_k^*)}{1-F'(z_k^*)} \zeta^c_{(z_k^*)} \Delta h_k^*(z_k^*) + \textstyle\int_{z_k^*}^{\infty} \tfrac{F'(z)}{1-F'(z)} \eta_{(z)} \Delta h_k^*(z) dz \quad - \int_{z_k^*}^{\infty} (1-g(z)) \Delta h_k(z) dz \quad - \int_{z_k^*}^{\infty} g(z) \hat{h}_k(z) dz \Big) \, .$$

Given that the sponsor is restricted to positive incentive schedules (i.e., $dz_k^* ds_k > 0$), the sponsor's expected surplus from the incentive schedule is proportional to

$$X_k \propto \underbrace{\left(\underbrace{\tfrac{z_k^* F'(z_k^*)}{1-F'(z_k^*)} \zeta^c_{(z_k^*)}}_{\text{S.E.} \geq 0} \Delta h_k^*(z_k^*) + \int_{z_k^*}^{\infty} \underbrace{\tfrac{F'(z)}{1-F'(z)} \eta_{(z)}}_{\text{I.E.} \leq 0} \Delta h_k^*(z) dz - \int_{z_k^*}^{\infty} \underbrace{(1-g(z))}_{\text{social cost}} \Delta h_k(z) dz \right)}_{\text{Expected welfare gain of perturbation to incentive schedule}} \tag{1}$$

$$- \underbrace{\left(\int_{z_k^*}^{\infty} \underbrace{g(z) \hat{h}_k(z)}_{\geq 0} dz \right)}_{\text{Expected social value of income transfer}} .$$

The first term between parentheses is the expected welfare gain of the sponsor's incentive schedule. Except for knife-edge cases, this term is positive for some signal.[11] Without loss of

[11] The unconditional effect of the incentive schedule on welfare is zero because we assumed the tax schedule is optimal, and the incentive schedule can be viewed as a perturbation to the tax schedule. The unconditional welfare effect of some perturbation is by definition equal to the welfare effect of that perturbation conditional on signal k times the probability of signal k summed over all possible signals. Given that probabilities are positive and the unconditional welfare effect is zero, a negative welfare effect conditional on some signal implies the welfare effect conditional on some other signal must be positive. Finally, given that optimal tax schedules depend on the distribution of earnings, the optimal tax schedule conditional on a signal is different from the unconditional schedule, except for special cases (e.g., when all marginal taxes are zero, i.e., $F'(z) = 0$, and no value to redistribution exists, i.e., $g(z) = 0$). Hence, some signal must exist for which the welfare effect of the perturbation is positive.

generality, I refer to the signal for which the welfare gain of the perturbation is the most positive as signal k. The intuition for the first term being positive is that the optimal tax schedule is different for the conditional distribution of earnings than for the unconditional distribution of earnings because the optimal tax schedules depend on the distribution of earnings. Hence, conditional on the signal, some welfare-enhancing perturbation to the original tax schedule is possible.

The first term also shows the intuition for the mechanism by which the incentive schedule increases welfare. The incentive schedule increases work incentives at earnings levels that are more likely conditional on the signal than unconditionally (i.e., at z_k^* where $\Delta h_k^*(z_k^*)$ is positive and large), especially if the conditional density at higher earnings levels is lower than the unconditional one (i.e., if $\Delta h_k^*(z_k)$ and $\Delta h_k(z_k)$ are more negative for $z > z_k^*$). The incentive schedule increases the positive fiscal externalities stemming from extra labor supply due to the incentive effect at earnings level z_k^* through the term labeled "S.E." while limiting the negative fiscal externalities from reductions in labor supply above earnings level z_k^* due to the income effect though the term labeled "I.E." and limiting the net social cost of transfers for earnings levels above z_k^* through the term labeled "social cost."

Unless income transfers for earnings above z^* are not socially valued (so, unless $g(z) = 0$ for $z \geq z^*$), the sponsor does not capture the full expected welfare gain. The sponsor does capture all the positive fiscal externalities from extra labor supply from the substitution effect at z_k^* and all the negative fiscal externalities from reduced labor supply due to the income effect at earnings levels above z_k^*. However, the sponsor bears the full cost of the transfers for earnings above z_k^*, rather than only the cost net of the social value of these transfers. Hence, the increase in the sponsor's expected surplus from the incentive schedule is the expected welfare gain minus the expected social value of the transfers for earnings above z_k^*.

The sponsor therefore only offers an incentive schedule if the welfare gain of the incentive schedule is sufficiently positive that it exceeds the expected social value of transfers to higher-earnings individuals. This condition is more likely to be met when the signal is stronger, because a stronger signal increases the expected welfare gain from perturbing the tax schedule conditional on the signal. This condition is also more likely to be met when the social value of transfers to higher earnings levels is lower. The social value of these transfers is lower when the social welfare function is more redistributive (i.e., when $g(z)$ is smaller at higher earnings levels). In the extreme case of a tax schedule that is optimal for a Rawlsian social welfare function, the social marginal value of consumption, $g(z)$, is zero for all earnings levels except for the very lowest level of earnings. Hence, the social value of transfers for higher earnings is zero, and the sponsor would implement an incentive schedule conditional on signal k that maximizes social welfare, independently of the strength of the signal.

The results above apply to an incentive schedule that is constructed as a marginal positive

perturbation to the tax schedule. They therefore show that if a signal is sufficiently strong, an individual-specific incentive schedule exists that makes the individual weakly better off and increases expected sponsor surplus.[12] It is immediate that any incentive schedule offered by the sponsor, even if the schedule is not marginal, must make the individual weakly better off; otherwise, a rational individual would not accept it. Similarly, it must increase the sponsor's expected surplus; otherwise, a risk-neutral surplus-maximizing sponsor would not offer it instead of the marginal schedule, which was shown above to increase the sponsor's surplus. Moreover, because the incentive schedule is nonnegative (and therefore requires an expected payment by the sponsor), the only way the sponsor's expected surplus from the incentive schedule can be positive is when the incentive schedule increases the individual's expected fiscal externality enough that it offsets the expected direct cost to the sponsor of the schedule. Hence, any incentive schedule the sponsor offers increases the expected fiscal externality of the sponsored individual. Finally, because any fiscal externality of the individual's labor-supply decisions flows to the sponsor and because no non-fiscal externalities exist, the increase in sponsor surplus combined with the sponsored individual becoming weakly better off implies social welfare must increase for any incentive schedule the sponsor decides to offer. These results are formalized as follows:

Proposition 3. *(Effects of a sufficiently strong signal on a single sponsor). If (i) the sponsored individual is rational, (ii) the sponsor maximizes its expected surplus, (iii) the sponsor is risk neutral, (iv) the cost of measuring fiscal externalities and administering sponsor fees is negligible, (v) no non-fiscal externalities exist, (vi) the tax system is set optimally for the unconditional ability distribution and has positive marginal tax rates for some earnings levels (i.e., $F'(z) > 0$ for some z), and (vii) the sponsor can offer only nonnegative incentives, then a sufficiently strong signal of individual ability can be constructed (i.e., $0 < \Delta h_k^*(z_k) < \infty$ is large enough for some earnings level z_k, where the marginal tax is positive, $F'(z_k) > 0$) such that a sponsor receiving this signal on an individual it sponsors offers this individual an individual-specific incentive schedule that (i) weakly increases the surplus of the individual, (ii) increases the expected surplus of the sponsor, (iii) increases expected social welfare, and (iv) increases the expected fiscal externality (defined as positive for positive externalities) exerted by the sponsored individual.*

In subsection 4.1, I introduce a design refinement that allows the sponsor to also capture the social value of transfers to individuals with earnings above z_k^*. In this design refinement, the spon-

[12] Although we don't know whether a sponsor will receive a sufficiently strong signal, a sufficiently strong signal clearly can exist. Case 1, described above, showed that for the extreme case of perfectly informative signals, an incentive schedule that increases sponsor surplus always exists. Hence, perfectly informative signals are sufficiently strong. Moreover, the signal does not have to be perfect to be sufficiently strong. Because $\Delta h_k^*(z_k^*)$ is a change in density, possible values for $\Delta h_k^*(z_k^*)$ have no upper bound. Thus, except for the knife-edge case in which $F'(z_k^*) = 0$, constructing a signal where $\Delta h_k^*(z_k^*)$ is sufficiently large, but less than infinite, relative to values of $\Delta h_k^*(z)$ for $z > z_k^*$ that equation (1) is positive is always possible.

sor's change in surplus is equal to the welfare effect of a marginal incentive schedule. Given that any informative signal allows for a welfare-enhancing perturbation to the tax schedule conditional on that signal, the sponsor offers an incentive schedule that increases expected welfare regardless of the strength of the signal. This result is presented in Proposition 5 below.

Equilibrium outcome when sponsors compete and innovate.

Given that sufficiently strong signals increase a sponsor's surplus, sponsors have an incentive to make innovations that allow them to observe such signals. Thus, signals are not exogenous, because sponsors can invest in making innovations that yield more informative signals. The resulting innovations reduce the efficiency cost associated with the taxation of earnings. Although the model described the information advantage of sponsors in terms of signals, in practice, sponsors may come up with other innovations that allow them to distinguish between individuals with different abilities, such as screening techniques (e.g., based on the type of work assignments). The signals in the model therefore can be seen as a specific case of a broader class of innovations that sponsors can make to alter labor-supply incentives in a way that reduces the efficiency cost of income redistribution.

As described in the main text, the fee that sponsors pay to the government for each individual they sponsor is set such that the expected surplus for the marginal sponsor is zero. This level of the fee ensures that those sponsors that have the highest expected surplus from sponsoring get to sponsor individuals. Thus, sponsors that have more successful innovations in estimating individual abilities will outcompete other sponsors in getting to sponsor individuals. Any rational individual matched to a sponsor is weakly better off than in the equilibrium without sponsors, because sponsors can only offer nonnegative incentive schedules. Thus, the individual is equally well off if they choose an earnings level for which the incentive payment is zero, and the individual is better off if they choose an earnings level for which the incentive payment is positive. Thus, relative to a setup without sponsors, and holding fixed the tax schedule, some individuals are equally well off and others are better off.

If the signal of the marginal sponsor is not strong enough to generate an expected increase in the sponsor's surplus, the sponsor fee remains the same as when no sponsor receives a signal. In this case, only inframarginal sponsors that receive a sufficiently strong signal receive a positive surplus. If at least one sponsor receives a sufficiently strong signal, social welfare increases because this sponsor's expected surplus increases, the other sponsors' expected surplus stays the same, individuals are weakly better off, and the government's budget is unaffected.

Moreover, if the marginal sponsor's signal is strong enough that it receives a positive expected surplus at the sponsor fee that would occur if no other sponsors would receive signals, the sponsor fee increases to keep the expected surplus of the marginal sponsor at zero. The expected surplus of

the inframarginal sponsors is positive. Because individuals are weakly better off, expected sponsor surplus weakly increases, and government revenue increases, welfare is higher than in the setup without sponsors. The government can, if it chooses, use the increase in revenue from higher sponsor fees to reduce the tax schedule such that all individuals are better off than in the setup without sponsors. These equilibrium results are summarized in the following proposition:

Proposition 4. (Equilibrium Outcome with Sponsor Innovation and Contracting) *If (i) the assumptions of Proposition 3 hold, (ii) at least one sponsor receives a sufficiently strong signal that it offers an incentive schedule, (iii) sponsors have free entry, and (iv) sponsors are small enough that competition between them drives the expected surplus of the marginal entrant to zero, then sponsor innovation and incentive schedules that sponsors offer result in an equilibrium that, relative to the outcome without sponsors, (i) makes each individual weakly better off, (ii) weakly increases each sponsor's expected surplus, (iii) increases social welfare, (iv) weakly increases the sponsor fee sponsors pay to the government, and (v) results in the marginal sponsor earning an expected surplus of zero. Moreover, if the marginal sponsor receives a sufficiently strong signal that it receives a positive expected surplus at the sponsor fee that would occur if no other sponsors received signals, then, in addition, (vi) the sponsor fee increases and (vii) the government could choose to use the increased sponsor-fee revenue to reduce the tax schedule so that all individuals are better off.*

4.1 Design refinement: Better incentives for sponsors with signals of individual ability.

Proposition 3 established that if a sponsor receives sufficiently informative signals of individuals' abilities, it can offer individual-specific incentive schedules that increase both social welfare and its own surplus. The proposition establishes that the schedules, if offered, improve social welfare, not that they maximize social welfare. Moreover, if signals are not sufficiently strong, the sponsor does not offer incentive schedules at all. In this subsection, I show how a refinement to the design mitigates the first of these drawbacks and eliminates the second. This refinement induces a sponsor to offer individual-specific incentive schedules whenever it receives informative signals of any strength. Moreover, the resulting incentive schedules maximize social welfare among allowable incentive schedules, where allowable incentive schedules are defined as those that can be considered marginal relative to the tax schedule.

The proof of Proposition 3 highlighted two factors that limited the sponsor's incentive and ability to offer welfare-maximizing incentive schedules. First, the sponsor's surplus from offering an incentive schedule was equal to the social welfare effect of the schedule minus the social value of transfers from the incentive schedule to those with earnings above z_k^* (see equation (1)). Second,

the sponsor could only offer positive incentive schedules. The refinement addresses both of these limiting factors.

The first element of the design refinement is that the government reimburses the sponsor for the social value of transfers to those with earnings above z_k^*, which equals $\int_{z_k^*}^{\infty} g(z) \left(\hat{h}_k(z) dz_k^* ds_k\right) dz$ for the incentive schedule considered in the previous subsection. To do so, the government would need to announce a schedule $g(z)$ that gives for any earnings level z the marginal social value of consumption for individuals at that earnings level relative to the marginal value of public funds. The term $\left(\hat{h}_k(z) dz_k^* ds_k\right)$ simply denotes the amount that the sponsor's incentive schedule, conditional on signal k, pays to individuals at earnings level z. Thus, for each dollar the sponsor pays to someone at earnings level z, the sponsor would receive $g(z)$ from the government. For marginal incentive schedules, the sponsor's objective now fully aligns with social welfare.

The second element of the refinement is to give sponsors the authority to impose negative incentive schedules (within limits, discussed below). This way, they can also offer welfare-enhancing incentive schedules to individuals for whom they receive a signal for which a negative perturbation of the tax schedule is optimal. When a sponsor imposes a negative incentive schedule, that is, one that extracts payments from a certain individual, the sponsor would need to pay the government the social value of transfers extracted. Without this provision, the sponsor would have too great an incentive to impose negative incentive schedules, because those schedules transfer resources from individuals to the sponsor.

With these two elements in place, a sponsor can create any non-linear incentive schedule by combining arbitrarily many positive or negative incentive schedules of the form considered in the previous subsection. As noted above, the sponsor's objective function fully aligns with social welfare as long as it uses marginal incentive schedules. Thus, within the class of marginal incentive schedules, the sponsor offers an incentive schedule that maximizes social welfare. The third element needed to make this refinement work is therefore something that limits the incentive schedules to ones that are small enough that they can be considered approximately marginal, but large enough that the welfare gains from the incentive schedules are meaningful. Such a limitation on the size of the incentive schedules also mitigates concerns about horizontal equity related to people at the same earnings level, about whom the sponsor received different signals, facing different incentive schedules, with some receiving incentive pay and others having to make payments. The optimal degree of the size limitation on incentive schedules lies beyond the scope of this paper; it depends on the importance of horizontal equity and on the speed by which sponsors' surplus from incentive schedules starts to deviate from social welfare as these schedules become larger.[13] This design refinement is formalized in the following proposition:

[13] A conservative starting point for the size limitation might be that negative incentive schedules are limited to 2% of after-tax earnings. That is, sponsors are not authorized to demand incentive payments from any individual that

Proposition 5. *If (i) the first five conditions of Proposition 3 hold, (ii) the government induces sponsors to internalize the social value of their incentive payments to individuals by imposing a fee payable by sponsors to the government of $g(z)$ times incentive payments received by sponsors from individuals at earnings level z and a subsidy from the government to sponsors of $g(z)$ times incentive payments made by sponsors to individuals at earnings level z, (iii) sponsors are given the authority to impose negative incentive payments, and (iv) the size of a sponsor's incentive schedules is limited such that they are marginal, then a sponsor's change in surplus from offering an incentive schedule is equal to the social welfare effect of that schedule. Moreover, under the additional condition (v) that sponsors receive an informative signal of individual ability, sponsors offer social-welfare-maximizing individual-specific incentive schedules, subject to the limitation imposed by (iv), for signals of any strength.*

Determining the optimal nonlinear tax in a decentralized fashion.

The government can use the incentive schedules offered by sponsors to determine the shape of the unconditional optimal nonlinear tax schedule. It can do so by setting next period's tax function such that the tax at earnings level z equals the current tax at that earnings level minus the average incentive payments paid to individuals at that earnings level. Given that many sponsors exist, a single sponsor's incentive schedule would have a negligible impact on the tax schedule, and the sponsor would set its schedule the same way as it would if taxes were not adjusted in response to incentive schedules. First consider the case in which sponsors receive no signals. Given that their surplus is equal to social welfare, they would offer an incentive schedule only if it improves social welfare. Thus, the combined schedule (consisting of the tax and incentive schedule) improves social welfare over the original schedule. The government then adopts this combined schedule as the tax schedule for the next period. This process continues until no further welfare improvements are possible with marginal perturbations to the tax schedule. At this point, the tax schedule is welfare maximizing for redistributive preferences embodied by the function $g(z)$. Thus, rather than trying to design its own optimal tax schedule, the government could instead make its redistributive preferences and its need for revenue known through the function $g(z)$ and find the optimal nonlinear tax schedule through an iterative process based on its observation of the average incentive schedules chosen by sponsors.

exceed 2% of the individual's net earnings. Moreover, the government would limit reimbursements to the sponsor of the social value of positive incentive payments to incentive payments that are less than 2% of the individual's net earnings. Sponsors would be allowed to provide positive incentive payments that exceed 2% of the individual's net earnings, but they would not get reimbursed for the part of the payment that exceeds the 2% limit. No reason exists to forbid positive incentive payments in excess of the 2% limit, because incentive payments that are not reimbursed are always welfare improving if the sponsor chooses to make them, as long as the reimbursements are marginal (for the same reason that incentives schedules in Proposition 3 are welfare improving).

Also in the case in which sponsors have signals of individual ability, adjusting the tax schedule iteratively using the same updating procedure would make sense, except that only incentive payments below the limit specified in condition (iv) of Proposition 5 would be taken into account in the updating. The updating would ensure the tax schedule is roughly optimal "on average." The resulting tax schedule would provide the sponsors with more bandwidth to deviate from it with individual-specific incentive schedules.

5 Social insurance

The role of sponsors in social insurance combines elements of their role in human-capital investments and their role in taxation. A re-interpretation of the human-capital-investment model applies to ex-ante decisions of the individual that affect the risk of suffering an adverse event, such as unemployment or disability, that is covered by social insurance. A version of the taxation model applies to the individual's labor-supply decisions conditional on an adverse outcome having occurred, for example, when the individual has become unemployed or disabled. Thus, formally, the models from Sections 3 and 4 also apply here, though they have a somewhat different interpretation. This section describes how these two formal models apply to social insurance.

Ex-ante moral hazard.

An individual can take a range of actions that affect the risk of an adverse event. For example, different jobs have different risks of injury or layoff. Thus, the vector $I \in \mathbb{R}^m$, which previously represented possible human-capital investments, now represents the set of possible actions that affect the probability and severity of an adverse event. Actions could include the type of job to take (affecting layoff risk and injury risk), precautions taken in one's personal life and in actions at work (affecting injury risk), attitude on the job, and acquisition of new skills (affecting layoff risk). Like in Section 3, the effects of these actions depend on personal characteristics, $\theta \in \mathbb{R}^n$, and they entail private net returns $R : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and fiscal externalities $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. The fiscal externalities include the government's expenditures on social insurance programs as a result of any claims the individual makes on those programs. Other fiscal externalities, such as through the tax-and-transfer system, of actions I are also included in F.

As was the case with human-capital investments, individuals generally do not take the socially optimal action I, because they do not take into account the fiscal externality of their action. Additional reasons for failing to take the socially optimal action could include a lack of knowledge about the socially optimal action, self-control problems, or liquidity constraints. Together, sponsors and individuals have the correct incentive to take the efficient action because they are jointly

the residual claimant of the effects of the action. Although attaining the efficient outcome may be impossible due to information asymmetries or contracting costs, the sponsor has an incentive to induce the individual to take a less inefficient action by using the same mechanisms that were discussed before: offering incentives, innovation, and information sharing. Because the model is formally the same as in Section 3, Propositions 1 and 2 also apply to actions affecting ex-ante moral hazard related to social insurance programs.

Ex-post moral hazard.

If an individual suffers an adverse event and receives payments from a social insurance program that protects against income loss, they face an effective tax schedule, consisting of the regular tax schedule plus the rules that govern the receipt of the social insurance payments. For example, someone receiving disability insurance payments can earn income below some limit and loses eligibility for disability insurance payments once earnings exceed that limit. The formal model from Section 4 now applies to the effective tax schedule that a person receiving social insurance payments faces, and the results from that section carry over to ex-post moral hazard.

Given that effective tax schedules that include social insurance payments often include notches or effective marginal rates exceeding 100%, whether they are set optimally for any social preferences is unclear. If they are not set optimally for any social preferences, sponsors can offer incentive schedules that are welfare improving and make both the sponsor and individual better off. Even if the effective schedule is set optimally conditional on the adverse advent but unconditional on any other signals about the individual, a sponsor receiving a sufficiently strong signal about the person's ability can achieve a welfare improvement (Proposition 3). Moreover, the equilibrium with many sponsors also results in a welfare improvement (Proposition 4).

For social insurance that pays out for medical or assisted living care, rather than for income loss, a reinterpretation of the formal model from Section 4 applies to ex-post moral hazard. In the re-interpretation, z is medical spending (rather than earnings), $1/w$ is the unobserved measure of the person's effective medical services received per unit of medical spending (rather than 1 over the wage), and l is effective medical services received (rather than labor supply).[14] The function $F(z)$ still represents the fiscal externality the individual creates at level z, with positive values of F representing positive fiscal externalities. However, $-F(z)$ is now interpreted as medical spending that is paid for by social insurance, rather than as minus the tax schedule. The social insurance program determines $F(z)$, but its expenditures on medical services for individual i are reimbursed

[14]To shoehorn the reinterpretation literally into the model of Section 4, we would have needed to define $-z$ as medical spending and $-l$ as effective medical services. Instead incorporating those minus signs into the budget constraint and the derivative of the utility function with respect to l is more intuitive. The budget constraint therefore is changed from $c = C_0 + z - F(z)$ to $c = C_0 - z - F(z)$, where C_0 is the consumption level the person would have had absent medical spending. The utility function is redefined from having $\partial u_i/\partial l < 0$ to $\partial u_i/\partial l > 0$.

by the sponsor of individual i. As was the case in Section 4, the sponsor can offer the individual positive incentive schedules over z. Given that the model is formally the same, the intuition and results from that section carry over. If the sponsor can perfectly observe individual-specific returns to medical spending $1/w$, the sponsor's incentive schedule attains the efficient level of medical spending and eliminates all ex-post moral hazard. In the more realistic case in which the sponsor observes signals of the individual's return to medical spending, the sponsor's incentive schedule improves social welfare only if the sponsor receives sufficiently strong signals (Proposition 3 for a single sponsor and Proposition 4 for the equilibrium outcome).

Broader applicability.

This framework naturally applies to means-tested transfer programs, where transfers depend both on earnings and observable characteristics, or tags, in the terminology of Akerlof (1978)[1]. Sponsors can give their sponsored individuals incentives that reduce the likelihood of them becoming eligible for the programs in the first place (as modeled in Section 3). If the sponsored individual becomes eligible, the sponsor can offer incentives to reduce the efficiency cost associated with program participation (as modeled in Section 4).

Social Security Old-Age Insurance also fits in this framework. This program can be viewed as a particular set of age-dependent tax-and-transfer schedules. Thus, the stochastic observable variable on which payouts are conditioned is age (rather than, e.g.,, unemployment or disability). Just like conditioning variables for other forms of insurance, reaching a given age is stochastic because one could die before reaching that age, but unlike for other forms of insurance, we would not refer to reaching a given age as an "adverse event."

Sponsors can also improve efficiency of insurance markets that are currently in the private sector, such as health insurance that is currently provided through employers. The reason is that these other forms of insurance may have fiscal externalities that are borne by the sponsor but not by current providers in insurance markets. By becoming a provider of such insurance, the sponsor internalizes these externalities and now has the correct incentives to offer an incentive schedule to the sponsored individual to take a more efficient action. For example, medical care received though private insurers prior to age 65 has a spillover effect on medical care utilization after age 65 (McWilliams, Meara, Zaslavsky, and Ayanian, 2007[2]). Medicare covers the expense of this medical care, but because this expense is a fiscal externality, the sponsor would have to reimburse the government for it. A sponsor providing health insurance before the age of 65 would correctly take this fiscal externality into account, whereas current private providers of insurance do not. Another example is that private insurance that reimburses physical therapy or mental health care would likely also affect future job performance, which has fiscal externalities through the tax system. A sponsor that directly provides these types of insurance would take into account the

spillovers onto future earnings.

References for Online Appendix

[1] George A. Akerlof. The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review*, 68(1):8–19, 1978.

[2] J. Michael McWilliams, Ellen Meara, Alan M. Zaslavsky, and John Z. Ayanian. Health of previously uninsured adults after acquiring Medicare coverage. *Journal of the American Medical Association*, 298(24):2886–2894, 2007.

[3] James Mirrlees. An exploration in the theory of optimum income taxation. *Oxford University Press*, 38(2):178–205, 1971.

[4] Emmanual Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229, 2001.