Optimal Income Taxation with Present Bias

Benjamin B. Lockwood*

January 14, 2016

In progress. For latest version, please see
`http://scholar.harvard.edu/bblockwood`.

Abstract

Work often requires up-front costs in exchange for delayed benefits—one must search for a job before becoming employed, paychecks are typically delayed by a few weeks, and a promotion may come only after months or years of extra effort—and mounting evidence documents present bias over labor supply in the face of such delays. This paper studies the implications of such present bias for the optimal income tax schedule. I derive expressions for optimal tax rates as a function of observable elasticities and present bias, conditional on income. Present bias lowers optimal marginal tax rates, with a larger effect when the elasticity of taxable income is high. If labor commitment contracts are feasible, tax rates depend on the residual uncorrected degree of present bias. Residual bias can arise either because workers are naive or because commitment contracts may be infeasible due to limited liability constraints which prevent firms from imposing fines on workers who quit. I calibrate the model using both existing estimates of present bias and a new estimate of residual present bias using subjective well-being trends following US welfare reforms in the 1990s. All evidence suggests bias is concentrated at low incomes. Numerical simulations show that for modest redistributive preferences, optimal marginal tax rates are substantially negative across low incomes, comparable to those under the Earned Income Tax Credit in the US. Yet the model also generates novel results about optimal tax timing, with implications for improving the schedule of EITC payments.

*Benjamin Lockwood, Harvard University, lockwood@fas.harvard.edu. I would like to thank Raj Chetty, Emmanuel Farhi, Nathaniel Hendren, David Laibson, and Matthew Weinzierl for extensive discussions, as well as Charles Nathanson, Stefanie Stantcheva, Glen Weyl, and seminar participants at Harvard for helpful comments. This paper was previously circulated under the title "Present Bias and the Optimal Taxation of Low Incomes".

1 Introduction

How should income tax policy account for the fact that people make mistakes? A growing body of research extends models of optimal taxation to allow for misoptimization of a general form, quantified by "misoptimization wedges". But estimating these wedges is challenging, since misoptimization cannot generally be inferred directly from observed choices. As a result, policy implications often require a deep understanding of the *way* in which people misoptimize. This paper focuses on a particularly robust, and well measured, form of misoptimization—present bias. I draw from the large literature on time inconsistency, wherein people grow more impatient over intertemporal tradeoffs as they grow near, and I adopt the common normative assumption that such impatience is a mistake (hence present "bias").

The motivation for focusing on present bias is threefold. First, recent evidence suggests present bias generates substantial labor supply distortions. Workers exert greater effort as payday approaches, impatient individuals take longer to find work, and students put off unpleasant tasks to the last minute.[1] Since work often entails such delayed benefits in practice—whether due to job search, coarse pay periods, or infrequent promotions—accounting for that reality is of first-order importance in understanding the tax policy implications of misoptimization generally. Second, a sizeable empirical literature documents the magnitude of the present bias, including some evidence about how bias varies across income levels—a key determinant of optimal tax design—making plausible empirical calibrations more feasible than for other less well estimated sources of bias. Third, the specificity of the benchmark model of present bias provides discipline. Adding that behavioral model to the standard model of optimal taxation generates testable empirical predictions which are consistent with otherwise puzzling patterns in the labor supply literature.

The contributions of this paper are both theoretical and empirical. In the theory section, I generalize the benchmark model of optimal income taxation to allow for present biased workers. I first assume that workers are naive about their bias, in which case the misoptimization wedge maps directly to their "structural" degree of present bias (conditional on income). I use optimal control theory to derive necessary and sufficient conditions for the optimal tax in this case, including a "negative at the bottom" result. I also allow for substantial multidimensional heterogeneity—

[1]See Kaur, Kremer and Mullainathan (2015), DellaVigna and Paserman (2005), and Augenblick, Niederle and Sprenger (2013), respectively, and Section 4 for an extensive discussion of existing empirical evidence.

including differences in compensation delays—with novel and surprising implications for the optimal *timing* of tax payments. In particular, if some individuals are paid up-front, then it is beneficial to *delay* the collection of taxes and the payment of work subsidies. This novel result has implications for improving the timing of EITC payments—a topic of substantial policy discussion.

The second part of the theory section considers the possibility that workers are sophisticated and can mitigate their bias by signing labor supply commitment contracts with employers. In this case, optimal taxes depend on the residual uncorrected degree of bias at each income. If contracts are subject to a limited liability constraint preventing payments from workers to firms, this residual bias may remain positive even for sophisticated workers, particularly at low incomes. This model generates novel positive predictions—a U-shaped pattern of elasticities of taxable income, and an inverse U-shaped pattern of income effects—which appear consistent with patterns in the empirical labor supply literature.

Empirically, I draw on a wide range of evidence to estimate the misoptimization wedge due to present bias, conditional on income. I first review existing experimental evidence and structural estimations to calibrate the underlying (structural) degree of present bias across incomes. I then present an analysis—new to this paper—which estimates *residual* bias using trends in subjective well-being after the 1990s welfare reforms in the US. All evidence suggests bias is concentrated at low incomes.

The implications of present bias for optimal income tax policy are striking. In the final sections of the paper, I present numerical calibrations of optimal taxation for a range of parameters and redistributive preferences. Most notably, when redistributive motives are modest, optimal marginal tax rates are negative across a substantial range of low incomes. This finding suggests that corrective motives may provide a novel rationale for marginal work subsidies, such as the US Earned Income Tax Credit (EITC). Such policies are common across countries and enjoy strong bipartisan support among policy makers, and among economists.[2] Yet under the conventional assumption that the goal of income taxation is to redistribute resources to low skilled individuals, such policies are suboptimal.[3] Although the existence of negative marginal tax rates need not imply their optimality,

[2] Among policy makers, for example, President Obama and Paul Ryan have both advocated expanding the EITC for households without children.

[3] Optimal income taxes generally feature a large guaranteed minimum income, combined with *positive* marginal tax rates at all levels of earnings. Some have argued that marginal work subsidies may be optimal if labor supply responses are concentrated on the extensive margin (i.e., if workers respond to a tax increase by exiting the labor force rather

this paper revisits this debate and shows that negative marginal tax rates may indeed be consistent with an optimal tax schedule.

To preview the quantitative implications of present bias, Figure 1 plots the schedule of optimal marginal tax rates in the baseline model economy with and without accounting for present bias. Results are plotted for a modest degree of inequality aversion (see Section 5 for details), which tends to reduce marginal tax rates. Even in this case, marginal tax rates at low incomes are quite high under the rational optimum, generating a sharp divergence between that policy and effective marginal tax rates in the US (plotted here for households with two children). Stronger redistributive preferences exacerbate this divergence. However after accounting for present bias, optimal marginal work subsidies can match or exceed those generated by the EITC.[4]

Relation to the literature

This paper relates to a number of subfields in the optimal taxation literature.

Optimal taxation with misoptimizing agents. Farhi and Gabaix (2015) give the topic a general treatment, characterizing results in optimal taxation in terms of "behavioral sufficient statistics" (misoptimization wedges), without regard to the source of misoptimization. In independent research, Gerritsen (2015) characterizes optimal income taxation with abstract misoptimization, and attempts to measure status quo labor supply misoptimization using survey data.[5] This paper takes an alternative (and complementary) approach. By focusing on a specific model of misoptimization, I am able to generate new testable hypotheses and concrete policy recommendations. In this respect, the present paper is more similar to Spinnewijn (2014), which calibrates optimal

than continuously reducing earnings)—see Diamond (1980) and Saez (2002). However numerical simulations suggest this rationale may be quantitatively limited—the benchmark simulations in Saez (2002), for example, calls for a marginal tax rate of 10% on the first $4,000 of earnings, and much higher rates thereafter. Moreover, recent research has shown that such models generate a "participation credit"—a fixed sum paid to all labor force participants—combined with positive marginal tax rates at all positive incomes (Jacquet et al., 2013).

[4]The income tax here represents the integrated tax and transfer schedule, including programs such as welfare and Food Stamps. Kaplow (2007) notes that the phaseouts of such programs could offset the negative statutory marginal tax rates of the EITC, yielding positive net marginal tax rates. However analyses by the Congressional Budget Office (2012) and the Center on Budget and Policy Priorities (2014), as well as Kotlikoff and Rapson (2009), find net marginal tax rates remain negative—generally between −10% and −25%, for low income households with children.

[5]Gerritsen (2015) also uses subjective well-being data to estimate status quo labor supply misoptimization, adopting the approach, pioneered by Di Tella et al. (2001), of regressing subjective happiness reports on covariates and using the ratio of coefficients to estimate a marginal rate of substitution. Although that paper lacks quasi-experimental variation, raising concerns about omitted variable bias, it too argues that labor is undersupplied at low incomes, consistent with the findings in Section 4.

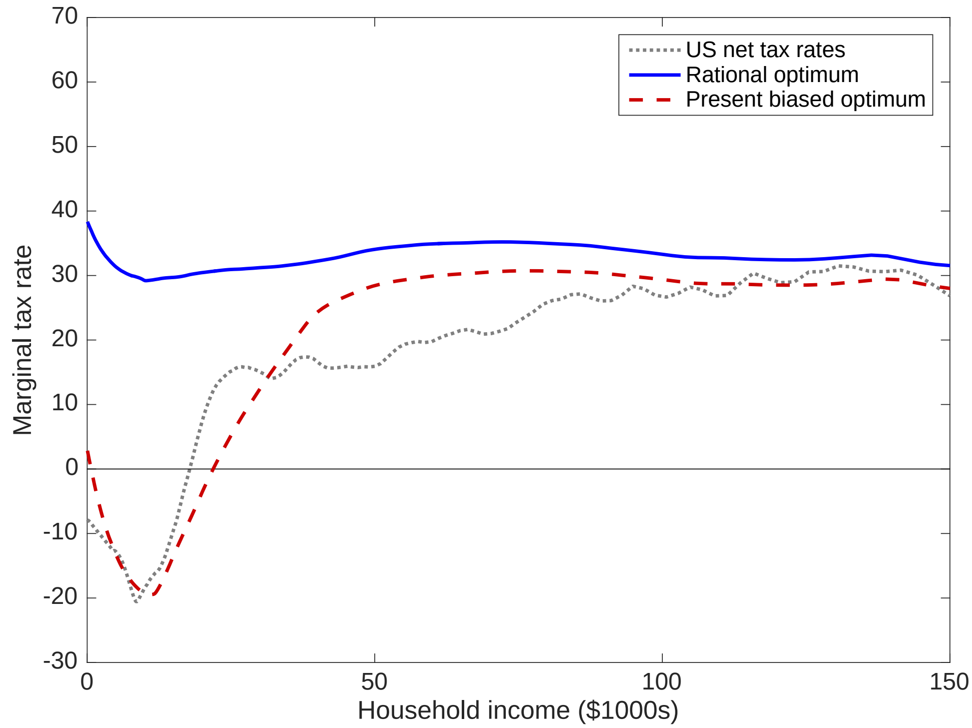

Figure 1: Simulated optimal marginal tax rates with and without present bias under modest redistributive preferences, compared to approximate net tax rates for a family with children in the US. The calibration of present bias is discussed in Section 4. The details of the model economy are discussed in Section 5, where the relationship between redistributive preferences and marginal tax rates is discussed at length.

unemployment insurance using data on mistaken beliefs about the probability of reemployment, and to Allcott and Taubinsky (forthcoming) and Lockwood and Taubinsky (2015), which perform quantitative calibrations of optimal taxation with consumers who misoptimize when consuming particular goods, such as those with delayed energy savings or uninternalized health consequences.

Optimality of negative marginal tax rates. As noted above, the reasoning in this paper presents a novel rationale for negative marginal tax rates. In the benchmark model of redistributive income taxation laid out by Vickrey (1945) and Mirrlees (1971), negative marginal tax rates are suboptimal (this finding has been discussed extensively—see Seade (1977), Seade (1982), Werning (2000), Hellwig (2007), and citations therein). A number of later analyses explored the sensitivity of this result to positive and normative assumptions in the conventional model. Diamond (1980) and Saez (2002) influentially argued that marginal tax rates could be negative at low incomes if earnings responses are concentrated on the extensive margin—for example, if there are heterogeneous fixed costs of labor force participation. Saez (2002) and Blundell and Shephard (2011) use models with discrete earnings levels to simulate optimal tax rates, finding low (or slightly negative) marginal tax rates on the lowest positive earning type. Jacquet et al. (2013) extended this insight to a continuous model, showing that extensive margin effects generally call for a positive participation credit (a fixed amount paid to all labor force participants), but that optimal marginal tax rates remain positive at all incomes—a structure quite different from the current EITC in the US, which features negative marginal tax rates on earnings up to $13,000.

Other work has shown that marginal work subsidies may be justified by normative objectives which differ from those in the conventional model. Most plainly, negative marginal tax rates may be warranted if the tax authority's goal is to redistribute income *upward*—i.e., if marginal social welfare weights (the social value of marginal consumption for households with a given level of earnings) are rising with income. This point was originally made in a discrete context by Stiglitz (1982); a number of recent papers have argued that such weights may arise from multidimensional heterogeneity, in particular preference heterogeneity (Cuff, 2000; Beaudry et al., 2009; Choné and Laroque, 2010).[6] Similarly, Fleurbaey and Maniquet (2006) show how fairness considerations may generate welfare weights which increase with income. Drenik and Perez-Truglia (2014) provide empirical evidence for such views, while noting that such an objective could generate Pareto inefficient policy recommendations. Although the reasoning in this paper is not inconsistent with such normative objectives, these findings demonstrate that negative marginal tax rates may be warranted even under the conventional assumption that policy makers wish to redistribute to low earners.[7]

The remainder of the paper is organized as follows. Section 2 presents a baseline model of optimal taxation with naive present biased workers and derives conditions characterizing the optimal income tax, including optimal tax rates and the optimal timing of tax collections, with a proof that

[6] Preference heterogeneity is not sufficient to generate negative marginal tax rates, however—see simulations in Lockwood and Weinzierl (2015), where optimal tax rates are lower in the presence of such heterogeneity, but remain positive.

[7] Two other proposed rationales for negative marginal tax rates have received somewhat less attention. First, negative rates may be justified by "non-welfarist" objectives, for example if the government wishes to minimize poverty (Kanbur, Keen and Tuomala, 1994; Besley and Coate, 1992, 1995) or has preferences directly over the labor and leisure choices of poor individuals (Moffitt, 2006). Second, work may have positive externalities—indeed, a more recent literature has demonstrated that labor force participation may have positive effects on children's outcomes (Dahl and Lochner, 2012; McGinn, Ruiz Castro and Lingo, 2015). This justification is not inconsistent with the present bias rationale in this paper—indeed the distinction between present bias and positive externalities on children is a blurry one. Although I will focus on shorter term benefits when calibrating present bias in Section 4, to the extent that these positive externalities are further undervalued, I will underestimate the size of optimal work subsidies.

delayed taxes can be optimal if some workers are paid up-front. Section 3 extends the environment to allow for sophisticated workers and commitment contracts. Section 4 estimates present bias conditional on income, drawing on several recent studies with a range of methodologies, as well as a new estimation of misoptimization using subjective well-being trends after the EITC expansion in the mid 1990s. Section 5 presents numerical simulations of the optimal policy in the baseline economy, and the resulting welfare gains from accounting for present bias. Section 6 performs an "inverse optimum" exercise to compute the implicit redistributive preferences which would rationalize existing policy under the conventional (perfect optimization) model and allowing for present bias. Section 7 discusses implications for policy design, as well as several important limitations. Section 8 concludes.

2 Optimal taxation with present bias: naive workers

The economy consists of a population of individuals of measure one, indexed by i and distributed according to measure $\mu(i)$. Utility is a separable function of consumption c, labor effort z/w (where z denotes pre-tax income, and $w_i > 0$ denotes ability), and fixed costs of work χ_i, where

$$U_i(c,z) = u(c) - v(z/w_i) - \chi_i \cdot 1\left\{z > 0\right\}, \tag{1}$$

I assume increasing, weakly concave utility of consumption ($u' > 0$ and $u'' \leq 0$), and decreasing and strictly concave utility of production ($v' > 0$ and $v'' > 0$), and I normalize $v(0) = 0$.

The policy maker's problem is to select the income tax function $T(z)$ which maximizes "social welfare"—a weighted sum of total utility, with individual-specific "Pareto weights" given by α_i, and I use $z(i)$ to denote i's earnings:[8]

$$\mathcal{W} = \int \alpha_i U_i(z(i) - T(z(i)), z(i))d\mu(i), \tag{2}$$

[8] In this constrained setting, (2) is isomorphic to a planner who maximizes a convex transformation of individual utility, $\int \Psi(U_i)d\mu(i)$, where the weights are set to equal the marginal social value of utility, $\alpha_i = \Phi'(U_i)$, evaluated at the optimum.

subject to a budget constraint

$$\int T(z(i))d\mu(i) - E \geq 0, \tag{3}$$

where E denotes an exogenous revenue requirement.

I then introduce a simple modification to this otherwise conventional setup: consumption occurs in the period following labor effort, and consumers are present biased—they discount utility in future periods by β_i, so that i's earnings satisfy

$$z(i) = \arg\max_{z} \left\{ -v(z/w_i) - \chi_i \cdot 1\left\{z > 0\right\} + \beta_i u(z - T(z)) \right\}. \tag{4}$$

Here β_i can be interpreted as in a β, δ model of quasi-hyperbolic discounting (Laibson, 1997), where the long-run discount factor δ has been normalized to 1.

2.1 Necessary and sufficient conditions for optimal taxes in a simple case

I first focus on an instructive simplified case, motivated by Diamond (1998), the first-order (necessary) condition for optimal marginal tax rates can be written in terms of model primitives and it is possible to specify *sufficient* conditions for this formula to represent the optimum.

Specifically, I impose the following restrictive assumptions:

1. Quasilinear utility: $u(c) = c$.
2. Heterogeneous ability distributed between $w_0 \geq 0$ and $w_1 < \infty$.
3. Homogeneous fixed costs: $\chi_i = \overline{\chi}$ for all i.
4. Constant α_i conditional on ability w_i.
5. Constant β_i conditional on ability w_i.

Notationally, I let $F(w)$ denote the distribution of ability, and I abuse notation to write Pareto weights and present bias as functions of ability w rather than type i: $\alpha(w)$ and $\beta(w)$. I let $e_\ell(w) = \frac{v'(z/w)w}{v''(z/w)z}$ denote the labor supply elasticity evaluated at the optimal choice of earnings for ability w, and $e_\beta(w) = \beta'(w)w/\beta(w)$, the elasticity of present bias with respect to ability. Then the following proposition characterizes the optimal income tax policy.

Proposition 1. *Under assumptions 1–4, the nonlinear income tax* $T(z)$ *which solves* (2)–(4) *must satisfy the following expression at all points of differentiability:*

$$\frac{T'(z(w))}{1-T'(z(w))} = \frac{1+1/e_\ell(w)+e_\beta(w)}{wf(w)} \int_w^{w_1} \left(1-\frac{\alpha(x)}{\lambda}\right) dF(x) - \left(\frac{\alpha(w)}{\lambda}\right)(1-\beta(w)),$$

where $\lambda = \int_0^\infty \alpha(w)dF(w)$. *Moreover, this expression is sufficient to characterize the optimum if the schedule of earnings which satisfies the resulting first-order condition,* $\frac{v'(z(w)/w)}{w} = \beta(w)(1 - T'(z(w))$, *is strictly increasing in* w.

Proof. See Appendix A. □

If instead this schedule would generate decreasing $z(w)$, then marginal tax rates are discontinuous—$T(z)$ is kinked—with a range of abilities bunching at the same level of earnings.

The expression in Proposition 1 is a generalization of Diamond (1998), which is identical except for the appearance of $e_\beta(w)$, and the final *corrective term*, $-\left(\frac{\alpha(w)}{\lambda}\right)(1-\beta(w))$. Intuitively, the former effect accounts for the fact that if bias is decreasing with ability (so $e_\beta(w) > 0$) high ability types are less tempted to imitate low ability agents, and thus marginal tax rates are optimally higher. The corrective term, which appears via the Hamiltonian optimization in the proof, can also be derived using economic intuition. Unlike in the conventional model, if individuals are biased then the "decision utility" marginal rate of substitution is not equal to the "experienced utility" marginal rate of substitution, and the welfare benefit from inducing a small increase in earnings is proportional to the ratio of the former to the latter. When misoptimization is due to present bias, this ratio is simply the share of marginal benefits from labor supply which are not internalized in decision utility when labor supply is set. Since a share β of delayed compensation is internalized, the remaining share $1-\beta$ quantifies misoptimization, and appears in the corrective term. This quantification motivates the following formal definition:

Definition 1. *The "misoptimization wedge" for an individual* i *earning* z *is equal to*

$$1 - \frac{v'(z/w_i)/w_i}{u'(z-T(z))(1-T'(z))}.$$

This wedge can loosely be thought of as the additional marginal tax rate which would induce an unbiased individual to choose z.

The corrective term has three notable features. First, the correction is negative when $\beta(w) < 1$—lower tax rates induce greater labor supply and thus help correct present bias—and it is larger in magnitude when $\beta(w)$ is far below 1. This overturns the familiar result that $T'(z) \geq 0$ for all z (Seade, 1977, 1982), since the corrective term outweigh preceding (nonnegative) term, generating negative marginal tax rates. Second, the term is weighted by $\frac{\alpha(w)}{\lambda}$, reflecting the planner's greater concern for correcting biases among individuals whose welfare is highly valued. Third, and perhaps least obviously, for a given Pareto weight and degree of bias, the corrective term has a constant effect on $\frac{T'(z)}{1-T'(z)}$, implying that the correction is larger when $T'(z)$ is low in the conventional (rational) model. Thus any feature of the standard model which reduces tax rates also inflates the size of the corrective term. For example, a higher elasticity of labor supply $e(w)$ tends to increase the size of the correction. This implies that if marginal tax rates are negative due to the corrective term, a higher elasticity of labor supply will make them more negative still. This result contrasts with other proposed rationales for negative marginal tax rates, where a greater labor supply elasticity tends to *reduce* the magnitude of negative rates. This relationship is explored in more quantitative detail in Section 5.

Proposition 1 also amends the "zero at the bottom" result from Seade (1977, 1982), as captured by the following corollary:

Corollary 1. *If* $\beta(w_0) < 1$ *and* $z(w_0) > 0$*, then* $T'(z(w_0)) < 0$.

In words, if all individuals work, and if the lowest ability workers are biased, then optimal tax rates are negative on the lowest earners. The proof follows immediately from evaluating the expression in 1 at w_0, which, since $z(w_0) > 0$, holds even at w_0. This result highlights the distinction between this setup and flexible marginal social welfare weights, as there is no schedule of Pareto weights $\alpha(w)$ which would justify $T'(z(w_0)) < 0$.

2.2 Multidimensional heterogeneity

I now relax the simplifying assumptions 1–4 to derive a more general expression for optimal marginal tax rates which permits income effects and multidimensional heterogeneity across w_i, χ_i, and β_i. This derivation employs the calculus of variations, as in as in Saez (2001), to express the first-order condition in terms of observables, which are endogenous to the tax system. Although more general, this approach does not permit a simple characterization of sufficient conditions, and thus I will focus on the necessary (first-order) condition and assume that it is also sufficient for optimal policy for the remainder of the paper. (I will, however, confirm that the resulting tax function generates a schedule of income which strictly increases with ability in the numerical simulations in Section 5.)

The necessary condition for the optimal income tax can be derived by considering a reform to the optimal tax code which slightly raises marginal tax rates by $d\tau$ in a small band ϵ around an income level z^*. Let $\mathcal{I}(z') = \{i | z(i) = z'\}$, representing the set of people with earnings z' at the optimum. In what follows, it will be useful to distinguish between endogenous parameters, such as the compensated elasticity of taxable income ε, for a given individual at the optimum, denoted $\varepsilon(i)$, from the parameter's average across individuals with a given income, denoted $\overline{\varepsilon}(z') = \frac{\int_{\mathcal{I}(z')} \varepsilon(i) d\mu(i)}{h(z')}$. For brevity the dependence of endogenous parameters on the tax code, though acknowledged, is notationally suppressed.

The relevant observables, defined loosely here and formally in Appendix B, are the income density $h(z)$, the compensated elasticity of taxable income $\varepsilon(i) = -\frac{dz(i)}{dT'(z(i))} \frac{1-T'(z(i))}{z(i)}$, the income effect $\eta(i) = -\frac{dz(i)}{dT(z(i))}(1 - T'(z(i)))$, and the participation elasticity $\rho(z) = -\frac{dh(z)}{dT(z)} \frac{(z-T(z))+T(0)}{h(z)}$. I follow Jacquet et al. (2013) in defining these responses to include any circularities due to the curvature of $T(z)$, which leads the marginal tax rate to change as earnings are locally adjusted. The effect also depends on the redistributive preferences of the planner (or, isomorphically, the perceived concavity of individual utility) which is encoded using standard *marginal social welfare weights*—written as a function of income—defined to be the marginal social welfare from additional consumption for an individual i earning $z(i)$:

$$g(i) = \frac{\alpha_i u'(z(i) - T(z(i)))}{\lambda}.$$

As before, λ denotes the marginal value of public funds.[9] If the α_i weights are constant across individuals, marginal social weights are the same for all individuals with the same income. This has the implication that the government does not seek to redistribute across individuals with a

[9] One feasible reform raises revenue by reducing all individuals' consumption so as to lower everyone's utility by a constant amount. Thus at the optimum the marginal value of public funds must equal $\left[\int \left(\frac{1}{\alpha_i u'(z(i)-T(z(i)))}\right) d\mu(i)\right]^{-1}$.

given income who have different combinations of present bias and ability—a plausible restriction given the potential difficulties of disentangling the the two sources of income variation. Thus more generally, even if the the α_i weights vary across individuals, it may be natural to assume that the weights are selected to yield constant marginal social welfare weights conditional on income at the optimum. Therefore in the derivations to follow, I invoke the following assumption, which turns out to simplify the expressions and render them more transparent.

Assumption 1. *Marginal social welfare weights $g(i)$ and $g_{ext}(i)$ are the same for all individuals with a given income.*

For notational consistency, I'll use $\overline{g}(z^*)$ to denote the marginal social welfare weight on individuals earning z^*, though in light of Assumption 1, this is simply equal to the common $g(i)$ of all individuals earning z^*.

The tax reform in question has a number of effects. First, it mechanically raises $d\tau\epsilon\int_{z^*}^{\infty}h(z)dz$ in funds from all workers with earnings above z^*, at a welfare loss of $d\tau\epsilon\int_{z^*}^{\infty}\overline{g}(z)h(z)dz$, for a combined mechanical effect of

$$dM = d\tau\epsilon\int_{z^*}^{\infty}(1-\overline{g}(z))h(z)dz.$$

Second, it induces a change in earnings of $dz^* = -d\tau\frac{\varepsilon(z^*)z^*}{1-T'(z^*)}$ among the $\epsilon h(z^*)$ individuals who experience a change in marginal tax rates. This alters tax revenues, resulting in a *fiscal externality substitution effect* of

$$dS_F = -d\tau\epsilon h(z^*)\varepsilon(z^*)z^*\left(\frac{T'(z^*)}{1-T'(z^*)}\right).$$

The change in earnings also has a direct effect on the welfare of z^*-earners, since they misoptimize due to present bias—a "welfare internality." The change in welfare from the local earnings adjustments due to the reform through the substitution effect is equal to

$$dS_W = -d\tau\epsilon h(z^*)z^*\mathbb{E}\left[\varepsilon(i)g(i)\left(1-\beta_i\right)|i\in\mathcal{I}(z^*)\right].$$

The formal derivation is provided in Appendix C. The economic intuition behind this expression is straightforward: as in the preceding subsection, only a fraction β_i of the benefits of work are internalized by the individual—thus the remaining share $1-\beta_i$ represents the welfare internality which is not taken into account. This expression can be simplified, and rendered more transparent, by invoking Assumption 1 to factor out the marginal social welfare weight, and by defining

$$\sigma_\varepsilon(z^*) = Cov\left[\frac{\varepsilon(i)}{\overline{\varepsilon}(z^*)}, \frac{1-\beta_i}{1-\overline{\beta}(z^*)}\middle| z(i)=z^*\right],$$

the covariance between the relative magnitude of elasticity and relative bias among z^*-earners. Then dS_W can be rewritten as

$$dS_W = -d\tau\epsilon h(z^*)z^*\overline{g}(z^*)\overline{\varepsilon}(z^*)\left(1-\overline{\beta}(z^*)\right)\left(1+\sigma_\varepsilon(z^*)\right). \tag{5}$$

The term $\sigma_\varepsilon(z^*)$ highlights the key role of heterogeneity. If more biased individuals earning z^* are also more responsive to the tax reform (high $\sigma_\varepsilon(z^*) > 0$) then the corrective term is inflated.[10] As shown formally in Appendix B, if $v(\cdot)$ has an isoelastic functional form, then the compensated elasticity of taxable income is constant across individuals with the same level of earnings, and thus the covariance is zero.

A third effect of the reform is an intensive margin change in earnings due to an income effect, equal to $dz = -d\tau\epsilon\frac{\eta(z)}{1-T'(z)}$, among individuals earning more than z^*. If leisure is a normal good, η is negative—individuals raise their earnings in response to their increased tax burden. The resulting *income effect fiscal externality* is

$$dI_F = -d\tau\epsilon\int_{z^*}^{\infty}\eta(z)\left(\frac{T'(z^*)}{1-T'(z^*)}\right)h(z)dz.$$

As with the substitution effect, the income effect also generates a welfare internality. I denote the covariance between the relative income effect and bias—analogous to $\sigma_\varepsilon(z^*)$ above—as

$$\sigma_\eta(z^*) = Cov\left[\frac{\eta(i)}{\overline{\eta}(z^*)}, \frac{1-\beta_i}{1-\overline{\beta}(z^*)}\middle| z(i)=z^*\right].$$

[10] This is analogous to the observation in Allcott and Taubinsky (forthcoming), in the domain of commodity taxes, that the relevant measure of bias for policy is the average bias weighted by the elasticity of demand response of marginal consumers.

which, again employing Assumption 1, can be written

$$dI_W = -d\tau\epsilon \int_{z^*}^{\infty} \overline{g}(z)\overline{\eta}(z)\left(1-\overline{\beta}(z)\right)(1+\sigma_\eta(z))\, h(z)dz.$$

As with the substitution effect, $\sigma_\eta(z)$ is equal to zero if $v(\cdot)$ has an isoelastic form.

Finally, the reform generates a change in labor force participation as some individuals with $z > z^*$ drop out of the labor force in response to the increase in taxes. These effects can be written in terms of the participation tax rate $\overline{T}(z) = \frac{T(z)-T(0)}{z}$. The measure of z-earners who leave the labor force is equal to $-d\tau\epsilon\frac{\rho(z)h(z)}{(1-\overline{T}(z))z}$, so the resulting fiscal externality is

$$dP_F = -d\tau\epsilon \int_{z^*}^{\infty} \rho(z)\left(\frac{\overline{T}(z)}{1-\overline{T}(z)}\right) h(z)dz.$$

To write the corresponding welfare internality in the case of this discrete change, it is useful to define the *extensive margin welfare weight*, the change in welfare (per dollar) from an increase in consumption equal to $z - T(z)$ for an unemployed individual i,

$$g_{ext}(i) = \frac{\alpha_i\left(u(z-T(z)) - u(-T(0))\right)}{z-T(z)+T(0)} \cdot \frac{1}{\lambda},$$

and to let $\mathcal{I}_{ext}(z^*)$ denote the set of individuals indifferent between earning z^* and exiting the labor force at the optimum.[11] Then, as shown in Appendix C, we have

$$dP_W = -d\tau\epsilon \int_{z^*}^{\infty} \rho(z)\overline{g}_{ext}(z)\left(1-\overline{\beta}_{ext}(z)\right) h(z)dz.$$

At the optimum this reform must generate no first-order increase in welfare, implying that $dM + dS_F + dS_W + dI_F + dI_W + dP_F + dP_W = 0$.

This assumption holds if the marginal social utility of consumption depends only on one's level of consumption. This is the case if the α weights are constant, or more generally, if the weights are assumed to be selected so that $g(i)$ is constant conditional on consumption at the optimum.

[11] Formally,

$$\mathcal{I}_{ext}(z^*) = \left\{ i | v'(z^*/w_i)/w_i = \beta_i u'(z^* - T(z^*)))(1-T'(z^*)) \cap -v(z^*/w) - \chi_i + \beta_i u(z^* - T(z^*)) = \beta_i u(-T(0)) \right\}.$$

(Thus this assumption need not rule out the special case of quasilinear utility of consumption with welfare weights declining with income.)

Then using the notation $\overline{\beta}_{ext}(z^*) = \mathbb{E}[\beta_i | i \in \mathcal{I}_{ext}(z^*)]$, denoting the average degree of present bias among those indifferent on the extensive margin, the optimal income tax is characterized implicitly by the following proposition.

Proposition 2. *At the optimum, if Assumption 1 holds, then the optimal income tax satisfies the following condition at all points of differentiability:*

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{h(z^*)\overline{\varepsilon}(z^*)z^*} \times$$
$$\left\{ \int_{z^*}^{\infty} (1-\overline{g}(z))h(z)dz - \int_{z^*}^{\infty} \overline{\eta}(z) \left(\frac{T'(z)}{1-T'(z)} + \underbrace{\overline{g}(z)\left(1-\overline{\beta}(z)\right)(1+\sigma_\eta(z))}_{I_W \text{ correction}} \right) h(z)dz \right.$$
$$\left. - \int_{z^*}^{\infty} \rho(z) \left(\frac{\overline{T}(z)}{1-\overline{T}(z)} + \underbrace{\overline{g}_{ext}(z)(1-\overline{\beta}_{ext}(z))}_{P_W \text{ correction}} \right) h(z)dz \right\} - \underbrace{\overline{g}(z^*)(1-\overline{\beta}(z^*))(1+\sigma_\varepsilon(z^*))}_{S_W \text{ correction}}. \quad (6)$$

This expression resembles the standard first-order condition for optimal tax rates with participation and income effects, with the addition of the terms involving β, which result from individual misoptimization. Although the parameters in (6) are endogenous to the tax code, they provide a transparent guide to the effect of present bias on tax rates.

The *substitution correction* (S_W) reduces marginal tax rates at z^* to induce z^*-earners to raise earnings, thereby correcting their tendency to underwork. The *participation correction* (P_W) reduces marginal tax rates at z^* to lower the level of taxes at higher earnings, which induces consumers to enter the labor force at those incomes. The *income effect correction* (I_W) works in the opposite direction—it tends to raise the optimal marginal tax rate (recall that η is nonpositive) since raising the level of taxes on higher earners induces them to beneficially increase their labor supply through an income effect. In the absence of participation and income effects, and if $\sigma_\varepsilon = 0$, this expression is equivalent to the simpler representation in Proposition 1.

Although this result provides guidance about the relationship between present bias and optimal tax rates, it also highlights the importance of estimating present bias *conditional on income*—a question taken up in Section 4.

2.3 Optimal tax timing

The two period model in Section 2.2 is useful for its simplicity, but since consumption and labor each occur in only a single period, it rules out some policy instruments, such as tailoring the *timing* of tax payments. This is an area of substantial policy interest, discussed more in Section 7.

To compare alternative timing regimes, this subsection considers an infinite horizon model in which the optimal steady state income tax is characterized by the same first-order condition as in Proposition 2. This setup generates two surprising insights. First, if all compensation is delayed, then tax delays are approximately irrelevant. (Irrelevance holds exactly under $\beta\delta$ quasi-hyperbolic discounting with $\delta = 1$, and otherwise holds approximately.) This finding contrasts with the common intuition that since present biased individuals discount the future, work subsidies should be paid as soon as possible for maximal corrective effect. Intuitively, taxes can subsidize earnings only at or after the time of compensation—but not before—implying that they are discounted by approximately the same amount regardless of their timing. Second, if some compensation is paid up-front, at the time of work, then it is beneficial to *delay* taxes. These results highlight the advantage of adopting a precise behavioral model for misoptimization, as they are not apparent from a generic misoptimization wedge.

Steady state model with delayed compensation

I begin with the case in which all compensation is delayed by a single period and taxes are levied at the time income is earned. Individuals are again indexed by i and choose earnings $z^t(i)$ each period t. As before, the planner selects a nonlinear income tax, and the problem is assumed to give rise to a steady state in which each individual i earns $z^*(i)$ each period, and the income tax $T(z)$ is constant over time. Individuals choose earnings to maximize the quasi-hyperbolically discounted stream of future period utilities, taking their own steady state labor supply in other periods as given, so that $z^*(i)$ satisfies

$$z^*(i) = \arg\max_z \left\{ U_i(z^*(i) - T(z^*(i)), z) + \beta_i \left[\delta U_i(z - T(z), z^*(i)) + \sum_{t=2}^{\infty} \delta^t U_i(z^*(i) - T(z^*(i)), z^*(i)) \right] \right\}.$$

Here δ represents the rational (exponential) discounting factor, as in the standard β, δ model (Laibson, 1997). Period length is selected to reflect the length of the "present", which receives full psychological weight—in light of Kaur et al. (2015), this might be a week or two.[12] Thus period length is sufficiently short that there is likely very little rational discounting, and so I will use the approximation $\delta = 1$ in what follows. Although the stream of future utility does not converge for exactly $\delta = 1$, the corresponding first-order condition remains well-defined:

$$v'(z^*(i)/w_i)/w_i = \beta_i u'(z^*(i) - T(z^*(i)))(1 - T'(z^*(i))), \tag{7}$$

and therefore we can define behavior as that which arises in the limit as $\delta \to 1$.

The government is assumed to maximize steady state period utility,

$$\mathcal{W} = \int \alpha_i U_i(z^*(i) - T(z^*(i)), z^*(i)) d\mu(i), \tag{8}$$

equivalent to maximizing the exponentially discounted stream of welfare in each period in the limit as $\delta \to 1$. I rule out government borrowing, so that the budget constraint requires that steady state taxes sum to zero.

Note that Equations (7) and (8) are identical to the individual first-order condition and the planner's objective function in the two period model from Section 2. Therefore letting ε, η, and ρ denote the compensated elasticity, income effect, and participation elasticity of steady state income with respect to steady state taxes (defined as in Appendix B) then Proposition 2 characterizes the optimal steady state tax policy without modification.

The benefits of delayed taxes

This setup permits tractably modeling taxes which are paid with a delay. Suppose first that all compensation is delayed by one period. Notationally, let $T_j(z)$ denote a tax that is levied with a delay of j periods. Then we can immediately verify the following proposition.

Proposition 3. *If all compensation is delayed and if $\delta = 1$, then the optimal income tax $T_j(z)$, and the resulting level of social welfare, is independent of j.*

[12] This model can also be written to allow for more general time inconsistent discounting functions, although the key insights remain the same.

Proof. The first-order condition for optimization under a tax with a delay of j periods is

$$v'(z^*(i)/w_i)/w_i = \beta_i \delta u'(z^*(i) - T(z^*(i)))(1 - \delta^j T'(z^*(i))).$$

Invoking the assumption that $\delta = 1$, this equation is identical to (7), and thus the problem has the same solution and resulting social welfare, regardless of j. $\square$

This result should be understood as a limiting case, as there are likely to be other (unmodeled) costs of levying taxes with long delays, such as administrative or enforcement costs. In that case, this proposition should be understood to show that, contrary to a common intuition, the costs of delayed taxes are small and are not driven by present bias.

Now consider the possibility that compensation delays are heterogeneous, and that the tax authority, who observes only compensation and not effort, cannot distinguish between income paid with with different delays. (For simplicity, suppose delays vary across individuals, and are outside their control.) By the logic of the preceding proof, such heterogeneity is irrelevant if all compensation is delayed by *at least* one period: delayed compensation is discounted by β_i regardless of the length of delay.

However if some individuals are paid up-front (a delay of zero), then their behavior differs from that of other individuals with delayed compensation, with implications for optimal taxes, in two respects. First, their labor supply is unbiased, so the government has no desire to encourage greater labor supply. However the planner is willing to tolerate some oversupply of labor among these unbiased individuals (beginning from the Mirrlees optimum, the resulting welfare costs are intially second order) in order to achieve the benefits of correcting bias among workers with delayed compensation. Second, unlike workers with delayed compensation, present biased workers with up-front payments *are* sensitive to tax delays. Taxes levied up front are "in the present," and hence receive full weight, whereas delayed taxes (or subsidies) occur "in the future" and thus are discounted by β_i, muting their effect on labor supply response. Thus by employing delayed taxes, the planner can dampen the inefficient overprovision of labor among unbiased workers who are paid up-front. In the notation of Proposition (2), delayed taxes induce a positive covariance between bias and the elasticity of taxable income ($\sigma_\varepsilon > 0$). Intuitively, delayed taxes are a more targeted corrective instrument, and the size of the corrective term is magified.

To formalize this result, suppose that a fraction ϕ of taxes are collected (or paid) in the period following the corresponding income, with the remainder $1-\phi$ paid at the time of compensation. Then we have the following proposition.

Proposition 4. *Starting from the optimal contemporaneous tax, if some present biased individuals are compensated at the time of work, social welfare is strictly rising with ϕ.*

Proof. Begin with the optimal $T_0(z)$, and consider the steady state effect on social welfare of a tax reform in which a fraction ϕ of taxes are levied with a delay of one period. Let π denote the fraction of individuals receiving up-front compensation. By the logic of the proof for Proposition 3, the behavior of the $1-\pi$ individuals with delayed compensation is insensitive to ϕ. The first-order condition for steady state earnings $z^*(i)$ for an individual who is compensated up-front is

$$\begin{aligned} v'(z^*/w_i)/w_i &= u'(z^* - T(z^*))(1 - (1-\phi)T'(z^*) - \phi\beta_i T'(z^*)) \\ &= u'(z^* - T(z^*))[1 - T'(z^*)(1 - \phi + \phi\beta_i)], \end{aligned}$$

where I have suppressed the dependence of z^* on i for brevity. The welfare effect of i's change in earnings $\frac{dz(i)}{d\phi}$ is

$$\frac{dz^*(i)}{d\phi}\left[\frac{u'(z^* - T(z^*))(1 - T'(z^*)) - v'(z^*/w_i)/w_i}{\lambda} + T'(z^*)\right] = \frac{dz^*(i)}{d\phi}T'(z^*)\left[1 - \phi g(i)(1-\beta_i)\right].$$

Moreover, for sensible elasticities and income effects, $\frac{dz^*(i)}{d\phi}$ has the same sign as $T'(z^*(i))$—an increase in ϕ mutes the effect of taxes on earnings behavior, raising incomes if taxes are positive and lowering them if taxes are negative. Therefore the welfare effect from a change in i's earnings has the same sign as $1 - \phi g(i)(1-\beta_i)$. Thus when $\phi = 0$, this effect is strictly positive. □

The economic intuition behind this result is straightforward: beginning from optimal contemporaneous taxes, a perturbation toward delayed taxes has no first-order effect on welfare (since workers with up-front compensation are unbiased) but has a positive first-order effect on the budget constraint—reducing earnings for those receiving marginal subsidies and raising it for those paying positive marginal tax rates—and thus the reform is beneficial.

More generally, if the government chooses discretely between contemporaneous taxes and a delayed tax, we have the following corollary:

Corollary 2. *If $g(i)(1-\beta_i) < 1$ for all i, then a delayed tax is strictly preferable to a contemporaneous one.*

The proof follows directly from the preceding one, which shows implies that if $g(i)(1-\beta_i) < 1$, the the welfare change from raising ϕ remains positive even when $\phi = 1$, i.e. when all taxes are delayed.[13] This condition is easily satisfied under the schedules of welfare weights used in the simulations in Section 5, and those computed under the inverse optimum exercise in Section 6, for the range of estimates of β_i in Section 4.[14]

The desirability of delayed taxes is closely related to the role of the term σ_ε in Proposition 2. Intuitively, the planner favors tax instruments which induce a greater corrective response among individuals who biased. When some individuals are paid up-front, and are thus unbiased, work subsidies induce them to overwork inefficiently—a second-order cost (starting from the Mirrlees optimum with no present bias correction) that the planner tolerates in exchange for the first-order corrective benefits from subsidies for present biased workers. By employing delayed taxes, the planner reduces such inefficient overworking, effectively allowing for more targeted corrections. This both raises welfare and raises the size of the corrective term in the expression from Proposition 2, implying larger work subsidies (or tax rate reductions).

Finally, although this exposition has been discussed in terms of different individuals, the same logic could be extended to apply *within* individuals who face multiple sources of earnings.[15] Intuitively, delayed taxes allow the government to encourage effort on tasks with delayed benefits, which are pursued too little (due to present bias), while inducing less overprovision of effort on tasks with up-front payments.

[13] Note that this condition is a sufficient condition—and quite a conservative one—since it implies that the welfare change from raising ϕ remains positive at $\phi = 1$. However even if this change were negative, e.g., if the optimal ϕ lies between 0 and 1, then it is possible that $\phi = 1$ dominates $\phi = 0$, in which case the planner would prefer the delayed tax instrument.

[14] Since sections 5 and 6 have maximal marginal social welfare weights 1.2 or less, then $g(i)(1-\beta_i) < 1$ provided that $\beta_i > 0.2$—a condition which holds for all the estimates in Section 4.

[15] Formally, this extension requires two additional assumption in order to align exactly with the preceding proofs: quasilinear utility of consumption, and separable disutility of labor effort from different income sources.

3 Sophisticated workers and commitment contracts

I now allow for the possibility that workers are sophisticated and can sign private commitment contracts with firms to mitigate present bias. For simplicity, I use the finite-horizon model in this case, although the results extend to the infinite horizon model discussed in the previous subsection. Commitment commitments are modeled by allowing for a "period 0" during which an individual and firm sign an "employment contract" which maps individual production y to the compensation paid by the firm, $z(y)$. In period zero, the individual discounts both utility from consumption and disutility of labor by β, so there is no present bias wedge between the two.

In some cases, the ability to sign contracts eliminates the need for corrective taxation. For example, if individuals have full information about their labor leisure tradeoff and there are no restrictions on contracts, then this result holds trivially: the individual's "self 0" signs a contract promising the privately optimal level of labor effort, on penalty of sufficiently large punishment, and the misoptimization wedge is zero. The model in this section relaxes the assumption of perfect information in period zero and of an unrestricted contracting space, thereby demonstrating circumstances in which corrective taxation is optimal even in the presence of private contracting. Nevertheless, it must be stressed that the optimality of corrective taxation is considerably more fragile when individuals are sophisticated and contracting is feasible than in the baseline model from Section 2. As will be seen in this section, the addition of contracting does not alter the necessary condition for optimal tax rates in Proposition 2, provided that $1-\beta$ is understood to represent the degree of *uncorrected* present bias (i.e., the misoptimization wedge). Rather, these results alter value of the parameters therein, conditional on income—for example, the misoptimization wedge is zero at high incomes, and the elasticity of taxable income is non-constant.

3.1 Model with commitment contracts

Utility is defined as in 1. In period 0, individuals have the option to sign a contract mapping individual production y to compensation z—in which case they are bound by that contract in periods 1 and 2—or to forego a contract and work without commitment as in 2. Firms are assumed to exist in a competitive labor market, so that all surplus is captured by workers, and firms can commit fully to contracts. Firms face a "hiring cost" or "vacancy cost" κ from creating and filling a position with a worker. Such costs were irrelevant in the previous section, since the deduction of such costs from workers' compensation could be subsumed into the fixed costs of work χ without loss of generality. For consistency, however, in this section I will assume that even workers who do not commit, and who flexibly choose to generates a product of y in period 1, receive compensation of $z = y - \kappa$ in period 2.

I assume that ability w_i and present bias β_i are known to both the firm and worker in period 0, either because the relationship is repeated, or because the firm can verify the worker's reputation (costs of which contribute to κ). Fixed costs $\chi_i \in \left[\underline{\chi}, \overline{\chi}\right]$ are assumed to be subject to some uncertainty in period 0, which resolves in period 1. Let $F_i(\chi)$ denote the distribution of these costs, conditional on information in period 0, from which χ_i (private to i and unverifiable) is drawn in period 1; $F_i(\chi)$ is known to the worker and the firm.

Individual production $y(i)$ is selected by "self 1" and, if chosen freely, would be too low from the perspective of self 0. However self 0 can choose a labor commitment contract—which maps production y to compensation z—in order to alter self 1's incentives and potentially achieve a more desirable outcome. Thus the optimal contract maximizes the utility of the "principal", self 0, subject to a participation constraint imposed by firms (nonnegative profits), and an incentive compatibility constraint imposed by the "agent", self 1. Moreover, contracts are restricted by a "limited liability constraint", which requires that $z(y) \geq \underline{z}$ for all y. For simplicity, I will assume $\underline{z} = 0$, so that firms cannot impose fines on workers, no matter their level of labor supply.

Since the task at hand is to model the optimal contract for a given individual, dependencies on i are suppressed. For notational simplicity, let $\tilde{u}(z) = u(z - T(z))$ (embedding the tax function, which is taken as given), and let $\tilde{v}(y) = v(y/w)$. In what follows, I assume that $T(z)$ is convex or not "too concave", so that $\tilde{u}(z)$ is a concave function, and I assume that $\tilde{u}(0) > -\infty$. Formally, a contract can be written as a mechanism mapping realized fixed cost to production and compensation: $Y(\chi)$, $Z(\chi)$. The optimal contract for a given individual solves

$$\max_{Y,Z} \int_{\chi} \left[-\tilde{v}(Y(\chi)) - \chi \cdot 1\left\{Y(\chi) > 0\right\} + \tilde{u}(Z(\chi))\right] dF(\chi) \tag{9}$$

subject to self 1's incentive compatibility constraint

$$
\begin{aligned}
&- \tilde{v}(Y(\chi)) - \chi \cdot 1\left\{Y(\chi) > 0\right\} + \beta \tilde{u}(Z(\chi)) \geq \\
&\qquad\qquad - \tilde{v}(Y(\chi')) - \chi \cdot 1\left\{Y(\chi') > 0\right\} + \beta \tilde{u}(Z(\chi')) \quad \forall \chi, \chi', \quad (10)
\end{aligned}
$$

and subject to firm participation (IR) constraint of nonnegative profits,

$$\int_{\chi} (Y(\chi) - Z(\chi))\, dF(\chi) - \kappa \geq 0, \tag{11}$$

and subject to the limited liability constraint on contracts:

$$Z(\chi) \geq z_0 \quad \forall \chi. \tag{12}$$

The selection of the optimal contract can be simplified by way of the following lemma:

Lemma 1. *The optimal contract is characterized by a threshold $\chi^* \in \left[\underline{\chi}, \overline{\chi}\right]$, with $Y(\chi) = 0$ and $Z(\chi) = Z_0$ for all $\chi > \chi^*$, and with $Y(\chi) = Y^*$ and $Z(\chi) = Z^*$ for all $\chi < \chi^*$.*

Proof. See Appendix D. □

Employing Lemma 1, the selection of the optimal contract amounts to the optimal selection of the parameters Y^*, Z^*, Z_0, and χ^*. Thus the optimal contract can be rewritten:

$$\max_{Y^*, Z^*, Z_0, \chi^*} F(\chi^*) \left[-\tilde{v}(Y^*) - \mathbb{E}\left[\chi | \chi < \chi^*\right] + \tilde{u}(Z^*)\right] + (1 - F(\chi^*))\tilde{u}(Z_0) \tag{13}$$

subject to nonnegative firm profits,

$$F(\chi^*)(Y^* - Z^*) - (1 - F(\chi^*))Z_0 - \kappa \geq 0, \tag{14}$$

and to incentive compatibility for self 1,

$$-\tilde{v}(Y^*) - \chi^* + \beta \tilde{u}(Z^*) \geq \beta \tilde{u}(Z_0), \tag{15}$$

and limited liability,

$$Z_0 \geq z_0. \tag{16}$$

This setup gives rise to three distinct contracting regions, characterized by the following proposition:

Proposition 5. *For a given β and distribution $F(\chi)$, different levels of ability give rise to three distinct contract regions:*

1. *Full commitment. Above some ability level w^h, constraints* (15) *and* (16) *are both slack, and all individuals work. Everyone is paid $Y^* = Z^* - \kappa$, and the misoptimization wedge is zero.*

2. *No commitment. Below some ability level w^ℓ, no feasible contract is preferable to the outside option of working without commitment in period 1. In this case no contract is signed, and in period 1 individuals either work at self 1's preferred labor supply with a misoptimization wedge of $1-\beta$, or do not work at all.*

3. *Limited commitment. At intervening levels of ability, bias is partially corrected, and the misoptimization wedge lies between 0 and $1-\beta$.*

Proof. See Appendix E. □

3.2 A numerical example

To illustrate the dynamics of the model above, consider a simple numerical example with quasilinear utility of consumption, $u(c) = c$ and isoelastic disutility of labor effort necessary to generate earnings $v(y, w) = \frac{(y/w)^{(1+1/e)}}{1+1/e}$, where e is the elasticity of labor supply (with respect to a change in the tax keep rate) in a conventional model. As in Section 5, I set $\varepsilon = 0.3$, and I assume a flat tax with $T'(z) = 0.3$ for all z. I assume a population with heterogenous ability and present bias, with w and β distributed independently and uniformly over $[\underline{w}, \overline{w}] \times [\underline{\beta}, \overline{\beta}]$. (See appendix F for details of this example.)

Implications for labor supply behavior Before turning to the implications for optimal policy, it's worth noting two sharp predictions about the income response to tax reforms across the income distribution. First, the compensated elasticity of taxable income is U-shaped over the region where some individuals have limited commitment. Under the functional forms in this example, the income elasticity (compensated and uncompensated) is e for individuals with full commitment and no commitment, and it is zero for individuals with limited commitment. Second, the income effect exhibits an inverse U-shape over the income distribution. In this example, quasilinearity of utility from consumption implies that income effects are zero among individuals with full commitment and no commitment.

Figure 2 plots the average compensated elasticity of taxable income and the income effect as a function of income. The income effect, unlike the compensated elasticity, is not constant conditional on income, and the black line smoothes the simulated results using kernel regression. Mean income effects at each simulated income grid point are plotted by gray points in Figure 2. I also plot a kernel regression to smooth this relationship (with a bandwidth of 0.7), which illustrates the inverse U-shape relationship between the income effect and income.

While modeling labor supply in the presence of limited commitment contracts is not the central focus of this paper, these patterns are consistent with two puzzling patterns from the labor supply literature. First, some evidence suggests elasticities are U-shaped. For example, Chetty, Friedman and Saez (2013) finds an elasticity of 0.31 in the phase-in region of the EITC, and an elasticity of 0.14 in the phase-out region. Many estimates place the elasticity at higher incomes at rather higher values, suggesting a U-shape (see, e.g., Chetty (2012), which favors an overall value of 0.33). Second, Meyer (2010) notes the puzzling absence of a reduction in hours worked among single mothers in the phase-out region following the EITC expansions in the 1990s. Such individuals experienced both an increase in benefits and an increase in marginal tax rates, suggesting that income and substitution effects should reduce hours worked—yet hours appear to be stable or to have increased slightly among this group. However if the income effect is positive and the compensated elasticity is low, as suggested by the middle region of the the income distribution in Figure 2, then such a reform might generate no change in hours, or even a positive change. Of course, more work is required to determine whether limited commitment contracts are indeed responsible for these patterns, but it is perhaps noteworthy that this simple model generates predictions consistent with observed patterns which otherwise appear puzzling.

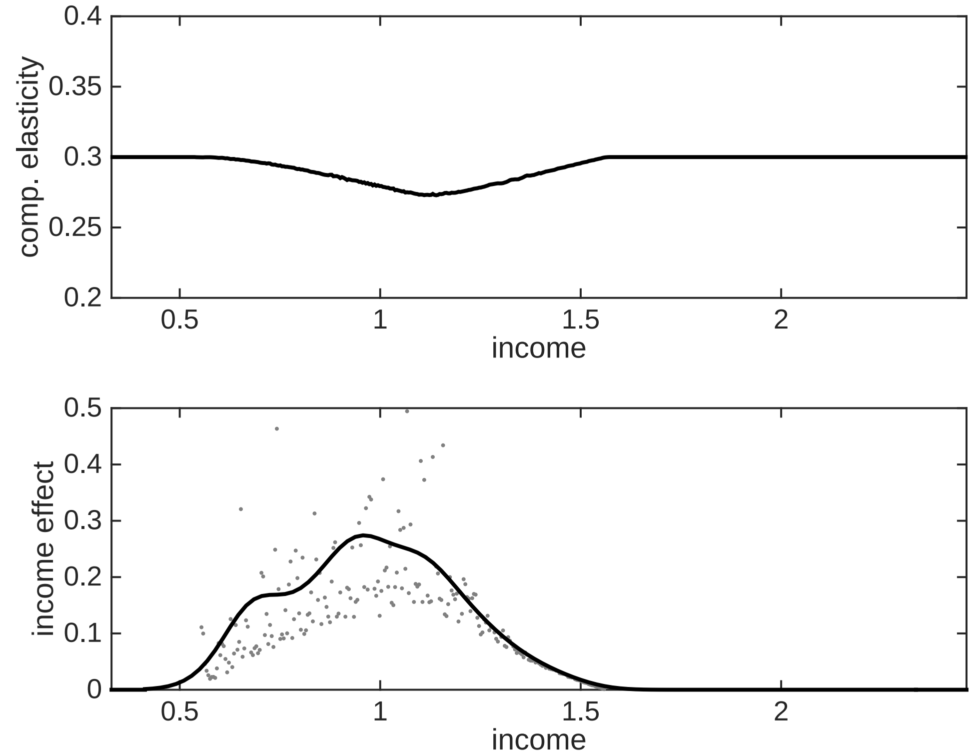

Figure 2: This figure displays the compensated elasticity of taxable income (top panel) and the income effect (bottom panel) as a function of income. Income effects are not continuous in income (individuals with similar incomes but different combinations of β and w have different income effects), so the bottom panel plots both the income effects in the simulated population (the gray points) as well as a kernel regression (the black line) to show the smoothed inverse U-shape pattern.

The misoptimization wedge. In the context of this model, the misoptimization wedge can be viewed as the residual or "uncorrected" present bias. For individuals in the "no commitment" region, the wedge is simply $1-\beta$. For individuals with full commitment, the wedge is zero. And for individuals with limited commitment, the wedge lies between 0 and $1-\beta$, reflecting the residual bias that remains after imposing the limited commitment contract.

Figure 3 plots the average misoptimization wedge conditional on income for this numerical example. Although the specific shape of the declining wedge in Figure 3 is the result of the assumed uniform joint distribution between ability and bias, the general shape—high at low incomes, declining to zero at high incomes—is a robust feature of this model.

This figure highlights a number of features which will prove relevant for policy design in the next section. First, the wedge is quite large at the bottom of the income, both because all such low earners lack labor commitments, and because within that group, individuals sort on income so

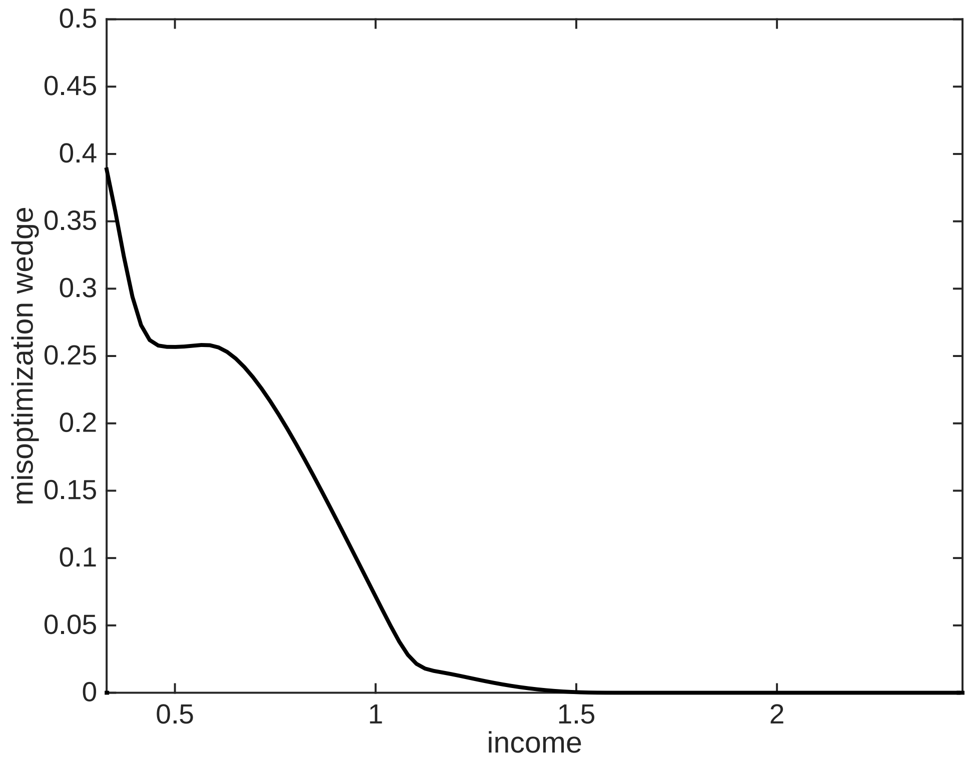

Figure 3: This figure displays the average misoptimization wedge conditional on income. The figure has again used a kernel regression to generate a smooth function of average bias by income.

those with the greatest bias earn the lowest incomes. At moderately low incomes, where all workers are still uncommitted, the bias plateaus at the mean level of $1-\beta$. The wedge then declines rapidly with income, as the share of workers with limited commitment and full commitment rises rapidly over this income range. Above the maximum income earned by workers without commitment, the wedge is quite small. This reflects the fact that although limited commitment is imperfect, most workers in the limited commitment range have largely mitigated their present bias, leaving little room for additional correction in this range. This suggests that the importance of the limited commitment individuals may be largely through their impact on elasticities and income effects, rather than for corrective policy. Finally, the wedge is equal to zero at higher incomes where all individuals are in the full commitment region. Thus in this region the familiar optimal taxation results, such as the first-order condition in Saez (2001) and the optimal top tax rate formula, apply without modification. In this sense, the implications of this affect policy design primarily at low incomes, where workers are marginally attached (or tempted to become marginally attached) to the labor force, even if substantial present bias exists among higher income individuals.

3.3 Implications of commitment contracts for optimal policy

Although the structural model in this section is more complicated than that discussed in Section 2, the expression for marginal tax rates given in Proposition 2 remains approximately correct, provided that the income-conditional bias wedge and income-conditional elasticities and income effects from Figures 2 and 3 are inserted in the formula. The reason this is an approximation is that the formula requires independence between behavioral responses and the bias wedge, conditional on income—although this is correct for the compensated elasticity, the income effect is larger for more biased individuals in the partial commitment contract region. (Intuitively, greater bias corresponds implies a greater need to adjust commitment to appease self 1's temptation to quit.) This inflates the benefit of targeting income to those who are partially committed. This effect is rather limited, however, by the small size of the misoptimization wedge in the limited commitment region (see Figure 3). Thus the expression in Proposition 2 remains a close approximation of the optimal tax in the face of commitment—and is exactly right at low incomes where no workers are bound by commitment.

Therefore the primary effect of incorporating private contracts is to account for the effect of commitment on the misoptimization wedge, and for the effect of contracts on the shape of elasticities. This highlights the importance of correctly measuring the "uncorrected" present bias conditional on income. In the next section, I present empirical evidence on the degree of structural present bias across income levels, as well as the extent of uncorrected residual bias relevant for the determination of optimal policy.

4 Calibrating present bias

Measuring the misoptimization wedge is a key challenge in generating policy recommendations from the model in the previous section. Unlike elasticities, misoptimization wedges cannot be estimated directly from responses to income tax reforms (succinctly: revealed preference does not identify misoptimization) so evidence on their magnitudes must be drawn from other, sometimes unconventional, sources.

Estimates of structural present bias					
			β estimate		
	low income/ EITC	labor supply	avg	low inc.	high inc.
Choice-based					
Augenblick, Niederle and Sprenger (2013)		✓	0.89	—	—
Kaur, Kremer and Mullainathan (2015)		✓	0.82	—	—
Jones and Mahajan (2015)	✓		0.34	—	—
Meier and Sprenger (2015)	✓		0.78	0.75	0.92
Structural					
Laibson, Repetto and Tobacman (2007)			0.70	—	—
Paserman (2008)		✓	0.65	0.40	0.89
Fang and Silverman (2009)	✓	✓	0.34	—	—
DellaVigna et al. (2015)		✓	0.59	—	—

Table 1: Estimates of present bias conditional from a variety of sources and methodologies. See text for details.

4.1 Estimates of structural present bias

Several recent studies quantify present bias. Table 1 reports estimates from a variety of such papers—details of their methodologies are discussed in Appendix G. I distinguish between two identification strategies: "choice-based" and "structural".[16] The choice-based approach identifies an environment in which choices are trustworthy (i.e., believed to reflect true utility) and assumes that similar preferences obtain in alternative environments where choices are believed to be suspect. In the context of present bias, choices over distant intertemporal tradeoffs are taken to be trustworthy; if they exhibit greater patience than choices over immediate tradeoffs, the difference identifies present bias.

Papers in the second category use structural models to estimate present bias, typically employing maximum likelihood methods to estimate a model with quasi-hyperbolic discounting. Further assuming that true utility is discounted time-consistently (exponentially), the estimate of β provides a structural estimate of the misoptimization wedge.

Table 1 also highlights papers with two desirable features. First, the column labeled "low

[16]These are two of the strategies for welfare analysis with behavioral models discussed in Chetty (2015); the third approach, direct measurement of utility, is employed below, in Section 4.2.

income/EITC" indicates papers whose subjects are low earners, and in some cases EITC recipients, in the US. Such studies are useful for two reasons. First, by they mitigate concerns about external validity, as their subjects resemble the population of interest for the simulations in this paper. Second, monetary rewards are known to be problematic for the estimation of present bias, since subjects can use their own funds to replicate (or undo) experimental variation in payoffs. However low income subjects are more likely to face liquidity constraints which prevent such arbitrage, perhaps explaining the substantial measured present bias even over monetary payoffs in those studies.

Second, the column labeled "labor supply" identifies papers which are estimated using intertemporal tradeoffs over labor supply, rather than money. These studies are particularly informative both because they avoid the shortcomings of monetary payoffs, and because the focus of this paper is labor supply, so to the extent that bias varies across domains, these studies identify the parameter of interest.

All studies find estimates of β meaningfully below one. Moreover, the two studies which permit the estimation of present bias across income levels—Meier and Sprenger (2015) and Paserman (2008)—suggest that bias is concentrated among low earners. (See Figure 5 in the next section for a graphical depiction of this relationship.) However the measures of present bias in Table 1 are not necessarily the ones relevant for policy. The choice-based and structural estimation methods in these studies identify the underlying ("structural") degree of present bias—β_i in the notation from Sections 2 and 3)—rather than residual bias after accounting for possible labor commitments in equilibrium. If workers are naive about their bias, as in the model from Section 2, then these measures are identical. However if workers are sophisticated and can sign commitment contracts, then the structural β_i will be lower than the residual bias.

4.2 Estimation of residual bias

This section presents an alternative method for measuring present bias—one which which identifies residual bias after commitment—using a strategy that measures a correlate of utility directly (subjective well-being) in response to an induced change in one's leisure-consumption bundle. Recall that residual bias is equal to the ratio of the decision utility marginal rate of substitution (MRS) to the experienced utility MRS. This approach draws from two sets of evidence to estimate each

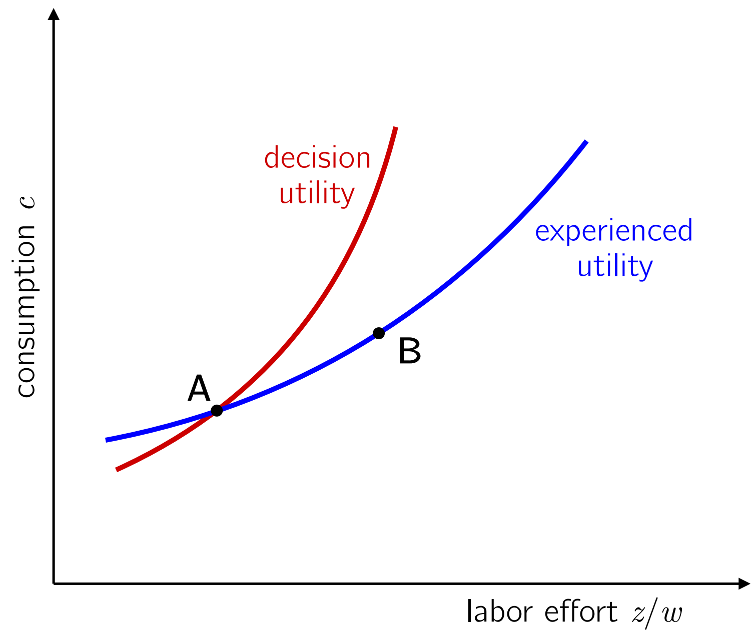

Figure 4: This figure illustrates the identification strategy used to estimate residual present bias. Points A and B denote the consumption-labor bundles of low income single mothers before and after the 1990s welfare reforms in the US. The decision utility MRS at point A is estimated using the net wage prior to the reform. The experienced utility MRS at point A is approximated by the slope of the line connecting points A and B, employing the fact that the data reject a decline in reported subjective well-being among low income single mothers over the reform period.

MRS separately.

The intuition behind this strategy is illustrated graphically in Figure 4. The figure plots the space of consumption-labor bundles, which has upward sloping indifference curves. When agents are present biased, decision utility undervalues consumption, and thus indifference curves are steeper for decision utility than for experienced utility. At a given consumption bundle, for example point A in the figure, the ratio of the slope of decision utility to the slope of experienced utility quantifies the misoptimization wedge, and is equal to $1/\beta$. The decision utility MRS can be measured using standard price theory logic: if A is freely chosen (subject to the individual's budget constraint), then the decision utility MRS at A is equal to the net wage, which is observable.

Measuring the experienced utility MRS is more difficult. Suppose the agent's budget constraint is altered by a tax reform, after which the agent is observed to choose point B, and suppose further that the individual's level of experienced utility was known to be the same at points A and B. Then by implication, A and B lie on the same experienced utility indifference curve, and the slope of the line connecting the points approximates the experienced utility MRS at point A. This slope can be computed as the ratio of the change in labor effort (hours worked) to the change in consumption.

This reasoning also illustrates why this method provides an estimate of *residual* bias, rather than structural bias. If point A is the result of a decision process that involves a commitment device which partially corrects structural bias, then the individual's "self 1 indifference curve" passing through A will be steeper even than the decision utility indifference curve plotted in the figure. Estimates of structural bias, as in the preceding section, measure the ratio of the slope of that steeper indifference curve to the experienced utility indifference curve, implying a greater degree of bias than is actually present after commitment. In contrast, this section estimates residual bias by using variation generated by the 1990s welfare reform in the US, with points A and B representing the consumption-labor bundles of low income single mothers before and after the reform. I perform the estimation separately for each of the bottom five income deciles of single mothers, effectively treating each decile as a single individual.

It's worth stressing several important limitations of this exercise. First, the assumption that the decision utility MRS equals the pre-reform net wage assumes that workers were at an interior optimum. If they were constrained from working as much as they desired prior to the reform, the net wage may overestimate decision utility MRS. The fact that Meyer and Sullivan (2008) find labor supply increased in response to the reform may provide some indication that higher labor supply choices were feasible. Although economic growth in the US during the late 1990s may have relaxed such constraints during that period, Meyer and Sullivan (2008) find similar labor supply increases when comparing the 1993–1995 period to either the 1997–2000 period or the 2001–2003 period, the latter of which should capture the ensuing economic contraction of the 2001 recession.

Second, this approach assumes that the consumption and labor supply measures examined by Meyer and Sullivan (2008) capture the variables relevant for experienced utility. This assumption is accurate if expenditures and hours worked are perfectly measured, if non-work time is spent enjoying leisure, and if agents live hand-to-mouth without saving—or, more precisely, if these potential confounds are not differentially active before vs. after the reforms. However if single mothers began saving more after the reform, for example, then total disposable income may have increased more than is indicated by the rise in consumption measured by Meyer and Sullivan (2008), biasing downward our estimate of the experienced utility MRS. Finally, this strategy is subject to the usual caveats associated with subjective well-being data.

At the same time, this estimation is a conservative estimate of bias in two important respects. First, the straight line connecting A and B is actually an overestimate of the experienced utility MRS at A, due to curvature of the indifference curve. Given the large measured change in consumption and hours worked due to the reform (about $1,500 and 550 hours, respectively) this curvature may not be trivial. Second, the calculation above assumes that experienced utility was known to be equal at points A and B. As will be seen below, evidence suggests that reported subjective well-being actually *increased* for low income single mothers after the reform.[17] This reasoning suggests the calculations below represent a lower bound for residual bias. On the other hand, the studies of subjective well-being trends (cited below) are not able to identify trends separately within each decile—further motivating the use of a conservative assumption on the change in well-being.[18]

In spite of these limitations, it is also worth emphasizing the general difficulty of measuring residual bias—yet since this parameter is ultimately relevant for policy, even coarse estimates provide useful information, and help test whether the estimates of structural bias appear to wildly different. Additionally, despite the notable limitations, the present exercise avoids some concerns present in previous attempts to quantify bias using subjective well-being data. In particular, Gerritsen (2015) estimates the experienced utility MRS without a source of quasi-experimental variation, and notes concerns about omitted variable bias.[19] Still, as subjective well-being data improves, employing this approach with larger data sets and across other natural experiments will no doubt generate more precise estimates of bias.

I now turn to the data sources used to implement this estimation strategy. The 1990s welfare reforms imposed time limits and restrictions on "lump sum"-like benefits, such as Aid for Families with Dependent Children, while augmenting work subsidies for single parents via an expanded EITC. Meyer and Sullivan (2008) examine the effects of the reform using data from the Consumer Expenditure Survey from 1993 to 2003, finding that both labor supply (annual hours worked)

[17]The data reject a decrease in single mothers' subjective well-being with $p < 0.01$ across all baseline specifications.

[18]Ifcher (2010) finds larger relative increases in reported subjective well-being when the sample is restricted to single mothers with only a high school education or below, suggesting the rises in SWB are unlikely to be concentrated especially at higher incomes within these groups. This provides some assurance that the operative assumption of an increase in SWB from point A to B is, if anything, especially likely to hold for lower income deciles, for whom residual bias is measured to be greatest.

[19]For example, if unobservable temporary depression spells reduce subjective well-being and one's willingness to work, it may appear that low labor supply *causes* lower happiness, spuriously suggesting misoptimization.

and consumption increased substantially for single mothers over this period. They compute the resulting implicit net wage by dividing the increase in consumption by the increase in hours spent working. The implicit net wages for the bottom five deciles of the consumption distribution among single mothers are \$2.11, \$2.69, \$3.21, \$4.67, and \$6.75 (in 2010 dollars). These figures correspond to the slopes of the lines connecting points A and B for each income decile. As the authors note, these findings suggest that if single mothers valued their time near their market wages (which were substantially higher), the reform made them worse off.

The second set of evidence is data on the evolution of subjective well-being among single mothers over this time period. Ifcher (2010) examines the evolution of reported subjective well-being among single mothers with children (who generally qualify for the EITC) relative to single childless women and to single men, before and after the substantial EITC expansion in 1996. The paper uses data from the General Social Survey, which since 1972 has asked respondents "Taken all together, how would you say things are these days–would you say that you are very happy, pretty happy, or not too happy?" Using a difference-in-differences analysis, the paper finds that self-reported happiness increased among single mothers after the welfare reforms of the 1990s, relative to both single women and single men without children. Consistent with these findings, Herbst (2013) uses data from a different survey with a shorter history (since 1985) but with a larger survey sample and state identifiers, and concludes that the reforms in the 1990s raised subjective well-being among single mothers. Boyd-Swan, Herbst, Ifcher and Zarghamee (2013) and Ifcher and Zarghamee (2014) also fail to find the decrease in reported happiness among single women that might be predicted from the findings in Meyer and Sullivan (2008).

Invoking the assumptions, noted above, that (1) single mothers working in the pre-reform period were at an interior optimum and (2) the subjective well-being evidence indicates that single mothers' happiness weakly increased within each consumption decile after the reforms, then the implicit net wages computed from Meyer and Sullivan (2008) represent upper bounds for the experienced utility MRS.

To compute the net wage of the bottom five consumption deciles of single mothers during the pre-reform period, I again turn to the CPS. To align with Meyer and Sullivan (2008), I use the sample of women respondents to the 1993–1995 March survey waves, restricted to women who are not married with a spouse present, who have at least one child, and who have positive wage income and total household income. I first construct measures of market income (total household income less welfare, unemployment insurance, social security and disability income) and net income (market income less federal income and payroll taxes net of credits, plus food stamps, and welfare income). I then divide the sample into deciles sorted by net income, to represent the deciles of single mothers (by consumption) in Meyer and Sullivan (2008), and I compute the net wage within each decile. The net wage is simply the individual's gross wage—equal to their self-reported wage, plus the employer portion of payroll taxes—multiplied by one minus the effective marginal tax rate. I estimate the effective net marginal tax rate by performing a 5th order polynomial regression of net income on market income, then predicting the local marginal tax rate at each level of market income. This approach appears to accord with the expected pattern of changes in net marginal tax rates due to the 1990s welfare reforms—the estimated net marginal tax rate on the bottom decile prior to the reform is 46%, whereas the same calculation on the analogous sample from the 2001–2003 March survey waves yields a negative net marginal tax rate.

The resulting estimates of bias are displayed in Figure 5, plotted along with the estimates of structural bias across income for the papers which permit such an estimation. As with estimates of structural present bias, residual bias appears to be heavily concentrated at low incomes. This is consistent with the prediction from the model with commitment in Section 3 that commitment contracts are most difficult to sustain among low earners, and thus that residual bias will be greatest in that region.

One distinction between the estimates of structural bias and residual bias is that the latter may encompass misoptimization for reasons other than present bias. If so, it may nevertheless be preferable to use estimates of residual bias, since the relevant statistic for optimal taxation is the *total* misoptimization wedge across all sources of bias. On the other hand, the similar shape of the schedules in Figure 5 suggests that present bias may be the primary source of bias relevant for labor supply misoptimization.

4.3 Implications for the misoptimization wedge

Figure 5 displays the resulting schedule of implied values of β for each of the data sources that identify bias across incomes. In each case, β is substantially below 1 at the bottom of the income distribution, and rises with income. The gray dotted line in Figure 5 shows the schedule of β across

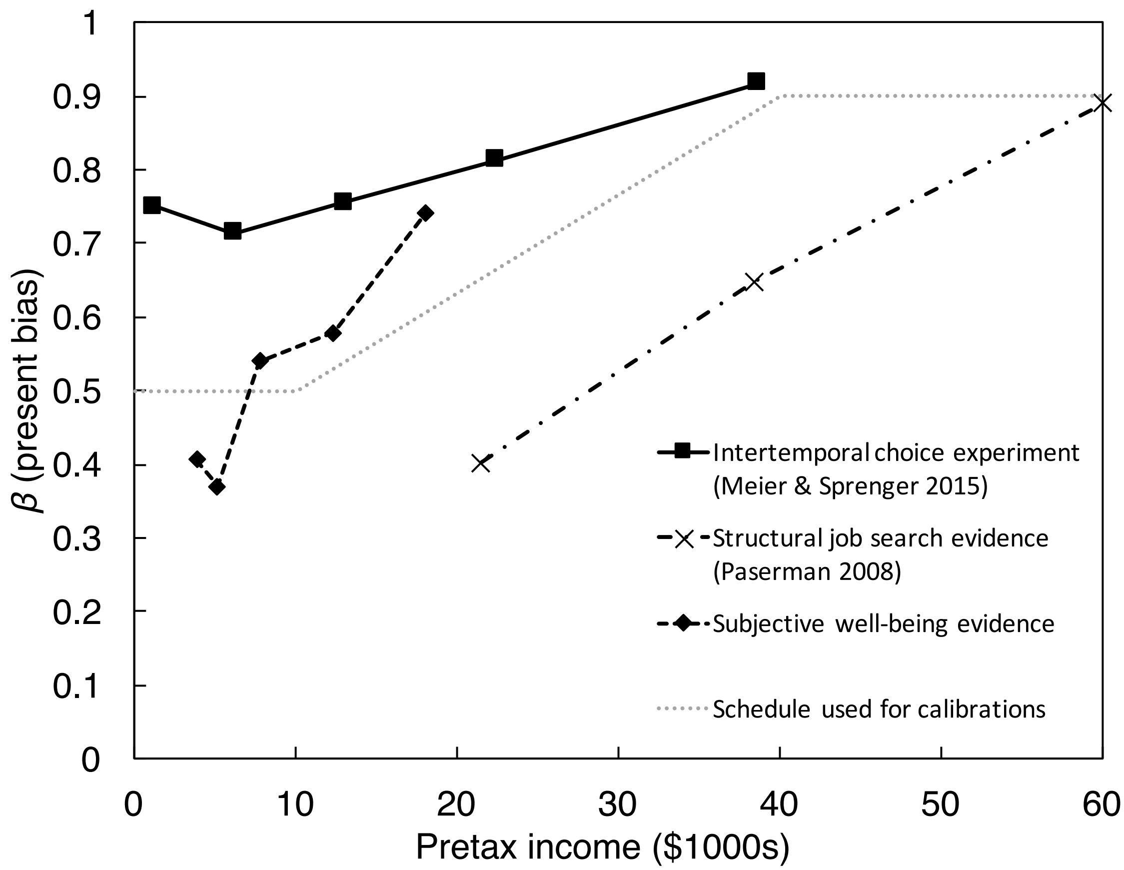

Figure 5: Estimated relationship between income and present bias parameter β from three sources (see text for details). The dotted line shows the approximation used for the simulations below.

incomes used for the simulations to follow. For those with annual market incomes below $10,000, I use a value of $\beta = 0.5$—approximately the average of the beta computed at the lowest incomes in each data source. For middle and high earners, I set $\beta = 0.9$. (The average value of β among the highest earners in each data source is 0.85.) The appropriate range of transition between $\beta = 0.5$ and $\beta = 0.9$ is rather uncertain—I interpolate linearly between market incomes of $10,000 and $40,000, as shown in Figure 5.

This relationship is consistent with a number of possible explanations. First, theory predicts that present-biased individuals endogenously select lower earnings. Second, present bias likely reduces human capital investments, leading to an inverse relationship between the bias wedge and underlying ability. Third, circumstances of material scarcity might *cause* greater present-bias (Mullainathan and Shafir, 2013; Shah, Mullainathan and Shafir, 2012). I remain agnostic about the mechanism for the relationship, effectively assuming that the plot in Figure 5 indicates a stable type-specific level of bias as a function of underlying ability.

5 Numerical analysis

In this section I present the details of the simulated economy used to generate the tax policy displayed in Figure 1. I also explore the effects of alternative normative objectives and modeling assumptions, where I show that the optimality of negative marginal tax rates like that in the EITC depends on the planner having fairly modest redistributive motives, especially across low incomes. In Subsection 5.3 I show that although the optimal policy accounting for present bias is quite different from that with rational agents, the welfare gains from accounting for bias are relatively modest—a dollar-equivalent gain of less than 1% of median income. A reform accounting for bias tends to benefit the "working class" (those between the 20th and 50th percentiles of the income distribution), while lowering the welfare of those in the bottom quintile.

5.1 The model economy

The primitives of the simulated economy consist of a specification of individual preferences, the policy maker's redistributive tastes, and the distribution of types (skills and biases). I assume individual utility is quasilinear in consumption, with $u(c) = c$, and disutility of labor supply is isoelastic, with $v(z/w) = \frac{(z/w)^{1+1/e}}{1+1/e}$. This form imposes two simplifications relative to the general form of individual utility in (1). First, it rules out income effects—a common assumption in the theoretical optimal tax literature (see, for example Diamond (1998) and Saez (2002)) which is consistent with empirical findings (Gruber and Saez, 2002; Saez et al., 2012). Second, it rules out discontinuous jumping from the extensive margin. I impose this restriction for two reasons. First, as shown by Jacquet et al. (2013), the addition of discontinuous jumping from the extensive margin does not generally rationalize negative marginal tax rates—rather it tends to generate a discontinuity in the tax function at zero, with marginal tax rates that look remarkably similar to the standard intensive margin model. So little is lost for the modeling of marginal tax rates, and some simplicity is gained, by omitting such discontinuities. Second, in practice many individuals whose responses are concentrated on the extensive margin tend to enter and exit the labor force multiple times per year. Indeed, data from the Current Population Survey indicate that even the lowest income decile of single mothers work an average of over 500 hours per year. When the extensive margin frequency is much higher than the tax period, it behaves more like an intensive margin of labor supply, and so I use an intensive margin as the benchmark model. I assume a fixed share of the population is disabled and has $w = 0$. To avoid complications of screening and optimal disability insurance, I assume that disability status is observable to the planner and that the required revenue for disabled individuals is exogenously given.[20]

I use a constant labor supply elasticity of $e = 0.3$—a value in the middle of empirical estimates (Chetty, 2012; Saez et al., 2012). Note that this parameter is similar to the compensated elasticity of taxable income from Section 2, however it this does not account for the curvature of the tax code. Thus the expression for optimal taxes in terms of this labor supply parameter would depend on the virtual income density (as discussed in Saez (2001)) rather than the actual income density as in Proposition 2.

I assume individuals are present biased according to the calibration from Figure 5. This provides a mapping from income to ability, via the individual's first-order condition for labor supply choice. Therefore I can infer the ability distribution from the income distribution from the Current Population Survey in 2010. Importantly, I infer the ability distribution taking account of the assumed degree of present bias, so that each policy represents the optimum given the *observed* income distribution and elasticities. I restrict to households with positive total family income (thereby excluding many with declared business and farm losses), and I use kernel density estimation to calibrate the density across incomes. The first-order condition depends on the individual's marginal tax rate, which is estimated from CPS and NBER's TAXSIM.[21]

The planner's policy objective is as in (2). Marginal social welfare weights are simply equal to the α-weights, normalized to have a mean of one. Note that this is isomorphic to the specification

[20] Specifically, I assume that 2% of individuals are disabled and unable to work altogether, consistent with the share of respondents in CPS between ages 25 and 55 with positive SSI income. I assume the exogenously required income for disabled individuals is $7,500, equal to average SSI income in this age group in CPS. Thus disability insurance effectively contributes to the government's exogenous revenue requirement.

[21] Specifically, I use TAXSIM's estimated net federal marginal tax rate, including employer and employee portions of payroll taxes, based on wage income, number of dependents, marital status, and age. I then construct an approximate implicit marginal tax rate from the phaseout of benefits using CPS data by performing a kernel regression of the value food stamps and welfare income on market income and differentiating the resulting schedule. This is also how the tax schedule in Figure 1 is computed, restricted to households with 2 children. I use a bandwidth of $2000 for the computation of marginal tax rates, and $5000 for the density estimation, where a greater degree of smoothing is useful for generating smooth schedules of simulated optimal tax rates.

of a concave transformation of individual utility, e.g. $\log(U(c,z,w))$, where the α-weights are set equal to the marginal utility of consumption at the optimum. For this reason, I'll adopt the conventional assumption that weights are declining as consumption increases. For the planner's budget constraint, I impose an exogenous government revenue requirement of $7,250 per capita, approximately equal to the net revenues currently raised by the federal income tax.

5.2 Optimal income taxes

This section demonstrates the key finding previewed in Figure 1: if redistributive tastes are modest, then negative marginal tax rates on par with those generated by the EITC may be approximately optimal.

For the baseline set of modest redistributive preferences, I select α-weights such that the lowest earners receive a weight of 10% more than the median household, and top earners weighted by 40% less than the median, linearly interpolated across intervening percentiles. These weights are substantially less redistributive than those conventionally assumed in the optimal taxation literature (e.g., logarithmic utility over consumption). However there is other evidence that the tax code embodies such modest redistributive tastes, e.g. through rather low tax rates at the top of the income distribution Lockwood and Weinzierl (2014).

Figure 6 displays the schedule of optimal marginal tax rates for four simulations. The top left panel is the baseline calibration reproduced in Figure 1. The top right panel shows optimal taxes with logarithmic redistributive preferences, so that $\alpha(w) = U(c,z,w)^{-1}$ at the optimum. The bottom two panels use baseline set of redistributive weights, with alternatively higher and lower labor supply elasticities.

These simulations highlight some key lessons for tax policy with present biased workers. First, as highlighted in the introduction, negative marginal tax rates like those under the EITC may be justified if redistributive preferences are modest. Second, as the log redistributive tastes case illustrates, if redistributive preferences are strong, marginal tax rates remain positive throughout the income distribution, although the inclusion of bias still tends to reduce tax rates below the optimum for unbiased individuals. This result may seem surprising in light of the fact that the bias term in Proposition 2 is weighted by the marginal social welfare weight, which is higher for lower earners under stronger redistributive preferences. That stronger corrective motive is outweighed,

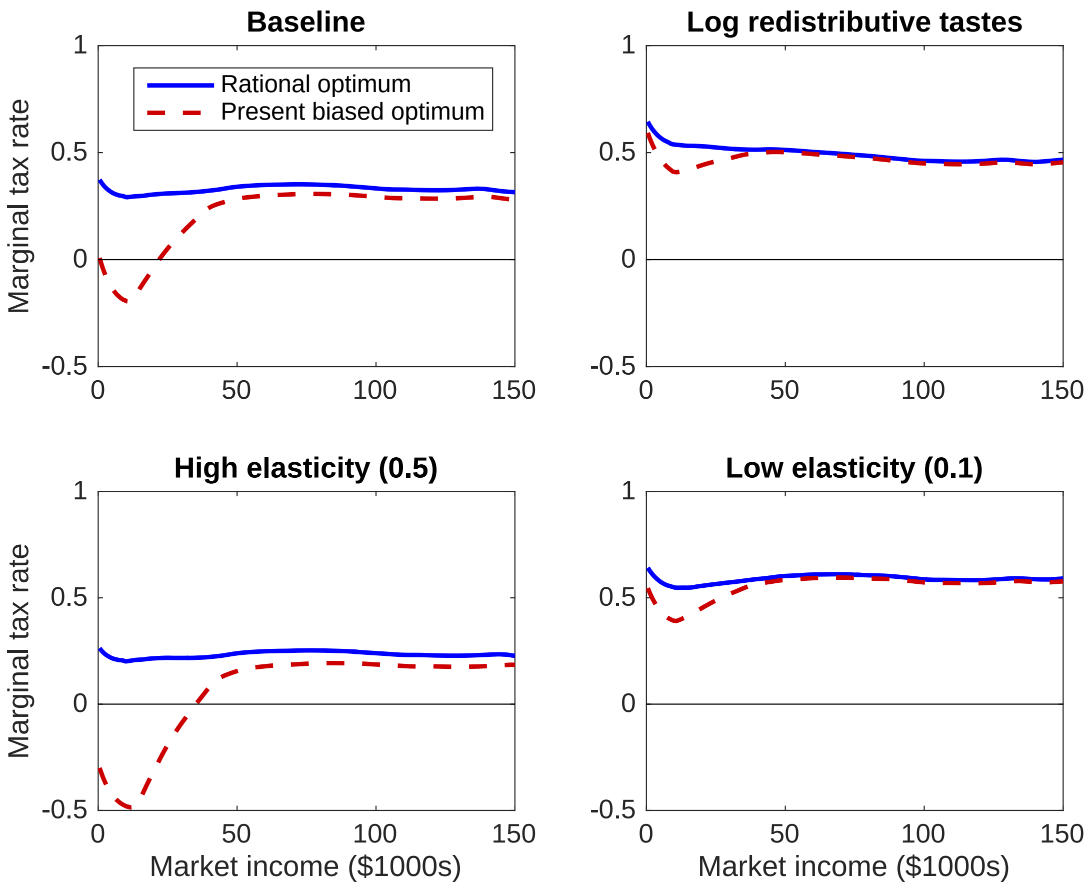

Figure 6: Optimal income taxes for alternative calibrations. The baseline uses a labor supply elasticity of 0.3 and modest redistributive preferences (bottom earners have a welfare weight 10% above the median, top earners have a welfare weight of 40% below the median, linearly interpolated). The second panel uses logarithmic redistributive preferences (see text for details), while the bottom two panels use the same modest redistributive preferences with alternative higher and lower values of labor supply elasticities.

however, by the stronger desire to redistribute *across* low earners under log preferences, reflected by the high level of marginal tax rates under the rational optimum in the log case. Since the corrective term operates on the term $\frac{T'}{1-T'}$, when marginal tax rates are high (close to one) in the rational case, this fraction is very large, and even a substantial corrective term has little effect.

The third lesson from Figure 6 relates to the labor supply elasticity. A higher elasticity has a strong effect on reducing optimal marginal tax rates in the present biased optimum. In fact an elasticity of 0.5, higher than baseline but still well within the range of some empirical estimates, particularly from the macro literature (see Chetty (2012)) generates optimal marginal tax rates

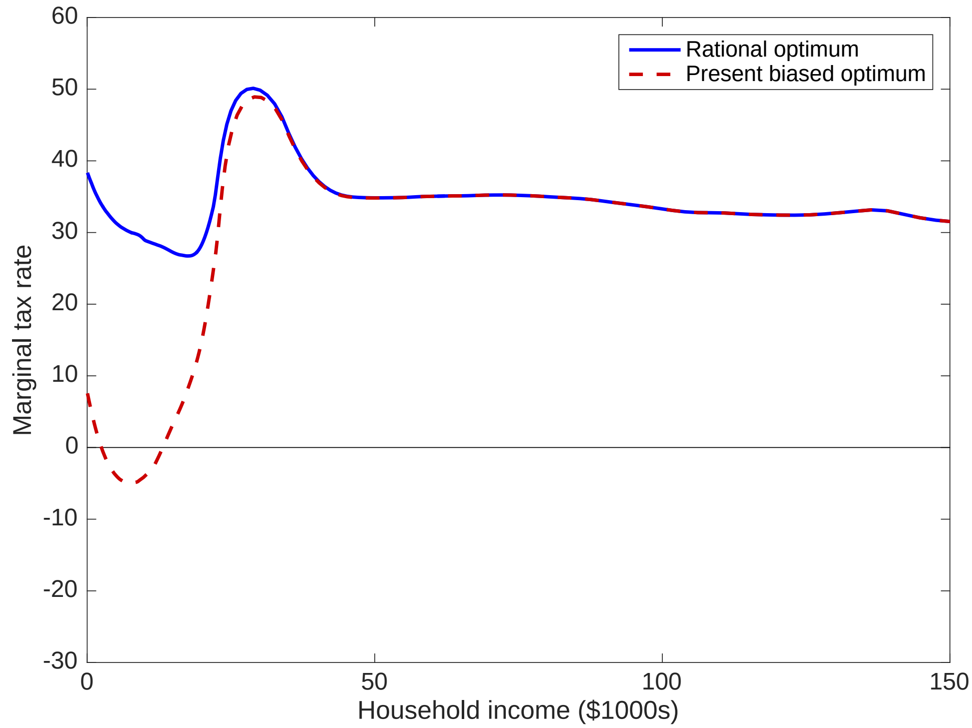

Figure 7: Approximation of optimal marginal tax rates with sophisticated agents and commitment contracts.

as low as -50%—subsidies much greater than the net rates generated by the EITC. It's worth stressing the divergence between this comparative static and the effect of higher intensive margin elasticities for the the EITC-like subsidies in Saez (2002), where a higher intensive margin labor supply *reduces* the magnitude of marginal tax rates on the poorest workers.

On the other hand, if elasticities are low, as in the lower right panel of Figure 6, then present bias does not generate marginal work subsidies at all. Of course, although these simulations use a constant value for the elasticity, the elasticity may vary with income in practice. Thus if individuals at the bottom of the distribution are particularly responsive to subsidies, they may be justified even if elasticities are fairly low at higher points in the income distribution.

To give a sense of the quantitative implications of contracting for the optimal income tax, Figure 7 modifies the simulation in the baseline economy in two ways. First, it is assumed that only a fraction of individuals are uncommitted at each level of income. I calibrate this share using the fraction of working individuals who work less than full time (2080 hours annually) in CPS. This is a coarse assumption, in the sense that some part time workers are likely committed to their privately optimal level of labor supply, while some who work 2080 hours or more may nevertheless be uncommitted along more flexible dimensions of effort (e.g., in exerting extra effort for a promotion). Still, this can represent a rough approximation of commitment in that many individuals who lose their jobs and spend some time searching for employment (thus, who are uncommitted at the margin) supply fewer than 2080 hours annually.[22] Unsurprisingly, in part due to the mechanical relationship between hours worked and income, the share uncommitted by this measure is much higher at low incomes (over 90%) than at high incomes, where it falls to around 30%. The second change is to assume a U-shaped pattern of elasticities—rather than a constant value of 0.3, I assume an elasticity of 0.3 for individuals with earnings below \$10,000 or above \$40,000 (under the status quo US income tax), with a U-shape between these incomes, falling to 0.2 at \$30,000, with spline smoothing in between.

It is worth stressing the illustrative nature of this calibration; it is ad-hoc in the sense that it assumes the share of workers and the elasticity is type-specific, rather than arising endogenously with a structural model of contracts. Nevertheless, the pattern of tax rates in Figure 7 is informative in two respects. First, marginal tax rates remain negative for a range of low incomes, although they are more muted than in the baseline specification in the baseline calibration. The reduction in the magnitude of the subsidy comes from the effective decrease in bias at low incomes, since a portion of low earners are now assumed to be setting labor supply correctly, and are in fact distorted by the corrective subsidies aimed at the biased individuals. Second, the U-shaped pattern of elasticities generates a hump in optimal marginal tax rates for the "working class"—those with incomes between \$25,000 and \$50,000. This suggests that the higher marginal tax rates in the phase-out region of the EITC—typically regarded as a necessary if undesirable feature of the credit by its proponents—may in fact be an optimal feature of the tax system.[23] This result is necessarily speculative in light of the substantial uncertainty about the correct model of firm and worker commitment contracts, but given the empirical evidence supporting a U-shaped pattern of elasticities, this possibility deserves greater exploration.

[22] Results look similar if I instead use individuals who work fewer than 52 weeks as a proxy for the share with uncommitted labor supply.

[23] As evident from the dotted line in Figure 7, these high phase-out tax rates are sufficiently diffuse across families with different incomes and tax situations that they do not appear in the smoothed net tax schedule.

Welfare gains from accounting for present bias		
	Dollars (per capita)	% of aggregate consumption
Baseline	$105	0.14%
Log redistributive preferences	$19	0.03%
High elasticity (0.5)	$243	0.3%
Low elasticity (0.1)	$12	0.02%

Table 2: Estimated welfare gains from implementing optimal tax, accounting for present bias, relative to a policy which appears optimal if the policy maker does not account for bias.

5.3 Welfare gains

This section considers the welfare gains from accounting for the existence of present bias—that is, for reforming from the "rational optimum" to the "present biased optimum" in Figure 6. The results are displayed in Table 2, which reports the welfare gains that result from reforming from the "rational optimum" to the "present biased optimum" in each of the specifications from Figure 6. The gains are measured both in dollars and as a share of the total gains from the optimal income tax.

Two features of these results stand out. First, the absolute size of the gains are modest—approximately $100 per household in the baseline specification. This finding is perhaps surprising in light of the large change in the tax code evident from Figure 6. One way to understand this "flatness" of welfare over the tax reform is through the reallocation of welfare across individuals. The reform to the optimal tax indeed raises efficiency by correcting the underworking of low earners, but to do so it sharply reduces the lump sum grant—from $19,500 to $6,900 in the baseline calibration. This reduces the welfare of the lowest earning individuals. However the low marginal tax rates mean that middle and upper-income individuals face a lower tax burden, and thus see their welfare rise. Thus while it is correct to say that in this framework marginal work subsidies "help" low earners overcome their bias by adding a corrective motive, when that motive is incorporated into the integrated optimum, it does not generate a net benefit for low earners relative to a policy which ignores present bias. (The implications of this conclusion for US policy may be rather limited, however, as the counterfactual optimum without present bias entails a far larger lump sum grant than exists in the US.)

The second notable feature of the welfare gains in Table 2 is the wide *variance* in the gains across calibrations—the dollar gain in the high elasticity calibration is over ten times larger than that in the low elasticity case. In general, welfare gains are larger when the elasticity of taxable income is high, and when redistributive motives across low earners are modest, since the efficiency gains from large marginal work subsidies on low incomes are particularly large in that case. This finding reinforces the importance, already prominent in the optimal taxation literature, of correctly measuring the elasticities at various points in the income distribution.

6 The "inverse optimum" approach and the EITC

The previous results have characterized the optimal tax schedule taking as given the redistributive preferences of the planner, through the marginal social welfare weights $g(z)$. Yet in reality there is great uncertainty about the appropriate set of redistributive preferences to apply for these calculations. One approach is to simply stipulate a schedule of marginal social welfare weights, with the caveat that the results are sensitive to them. This is the approach adopted, for example, by Saez (2001, 2002) and Jacquet et al. (2013).

In this section I adopt an alternative approach, inverting the optimal policy simulation by taking existing policy as given and computing the redistributive preferences with which those policies are consistent. This strategy, implemented by Bourguignon and Spadaro (2012) for European countries and further explored by Hendren (2014) and Lockwood and Weinzierl (2014), provides a reduced-form way to check whether existing policy generates weights which appear reasonable. Of course, the definition of "reasonable" is itself controversial in this context, but two features are commonly thought to be sensible requirements in the optimal policy literature: weights are positive at all incomes (Pareto efficiency) and weights are declining with income (redistribution toward lower abilities). In this section, I show that the implicit weights in the EITC population which arise under the usual assumption of perfect optimization exhibit a robust "unreasonable" feature: weights rise substantially with income across the bottom quarter of the income distribution. Taken at face value, the weights suggest that current policy implicitly places greater value on a marginal dollar for the median earner than on a dollar in the hands of the poorest EITC-receiving households, typically working single mothers. I compare these weights to those which arise implicitly from a model which allows for misoptimization due to present bias.

A secondary strength of the inversion approach is that it permits a more detailed representation of the complexities of the actual economy. Since this approach involves only an inversion of the local first-order condition for optimal taxes, it does not require a structural model of earnings responses to non-local tax reforms. This is particularly beneficial if individuals can sign labor commitment contracts as in Section 3, since a structural model of that commitment process may be infeasible. As a result, it is possible to incorporate a rather more detailed calibration of elasticities, including participation elasticities and non-constant compensated elasticities of taxable income. This section can therefore be viewed as a sort of robustness check for the result, from Section 5, that after accounting for present bias, the EITC is consistent with plausible redistributive weights, even under a more detailed calibration of the economy, whereas the redistributive motives implied if one does *not* account for bias appear rather unconventional.

6.1 Calibrating the income distribution and tax schedule

In the interest of computing the weights generated by the EITC, I focus on the set of families affected by the credit. Therefore I use a sample different from the one in the benchmark economy of Section 5, though I continue to draw data from the CPS. The income distribution is drawn from CPS, using the March survey waves for the years 2001–2010. (All figures are reported in 2010 dollars, adjusting using the CPI-U.) I restrict to nonfamily householders and to households for whom the respondent is the head of household, between the ages of 25 and 55, and I restrict to households with two children and with positive total income. I drop families with income below $30,000 who do not receive the EITC (23% of all households in that earnings range). A continuous income distribution is constructed by discretizing the income space into $1000 bins and using a fifth order polynomial regression on the number of households in each bin to generate a smooth density with a continuously differentiable derivative.

The schedule of marginal tax rates is drawn from the National Bureau of Economic Research's TAXSIM model. To compute the marginal tax rate at each point in the income distribution, I submit data on year, filing status, and the number and age of dependent children to TAXSIM, which provides an effective marginal tax rate on additional earnings, accounting for credits and deductions. I include the marginal tax rate from payroll taxes (both the employer and employee portions). I then average these marginal tax rates within each $1000 bin, and use these averages to compute marginal social welfare weights.

6.2 Parameter assumptions

In addition to the income distribution and schedule of marginal tax rates, the optimal tax condition depends on the compensated elasticity of taxable income, the income effect, and the participation elasticity, as well as the misoptimization wedge, conditional on income.

Beginning with compensated elasticities, I assume an elasticity of 0.3 at middle and high incomes—a central value in the range of existing estimates, and near the preferred value in Chetty (2012) of 0.33 to reconcile existing micro and macro estimates in the presence of optimization frictions. For the elasticities at low incomes, I draw from evidence drawn specifically from the EITC-receiving population. Chetty, Friedman and Saez (2013) estimate intensive margin elasticities of 0.31 and 0.14 in the phase-in and phase-out regions of the EITC, respectively, identified by differences in knowledge of (and, by assumption, responses to) the EITC across regions. Therefore I assume a compensated elasticity of 0.31 for households with incomes below $13,000 (the approximate upper bound of the phase-in region) and 0.14 for those with incomes between $13,000 and $40,000.

For the participation elasticity, Saez (2002) performs calibrations using estimates of 0, 0.1, and 0.5 for the bottom half of the population (and zero for the top half). Chetty et al. (2013) estimates an elasticity of 0.19 among EITC recipients. To avoid a discontinuous drop in the participation elasticity conditional on income, I interpolate (linearly) between $\varepsilon = 0.2$ at the bottom of the income distribution, declining to $\varepsilon = 0$ at an income of $40,000$. I further assume that income effects are zero throughout the distribution (Gruber and Saez, 2002; Saez et al., 2012).

As in Section 5, I assume the schedule income-conditional present bias is as reflected in Figure 5.

6.3 Implicit welfare weights on EITC recipients in the United States

The resulting marginal social welfare weights are plotted in Figure 8, both under the conventional assumption of no misoptimization, and under the assumption that individuals are present biased. In each case, the plotted points represent the weight computed locally for each percentile of the

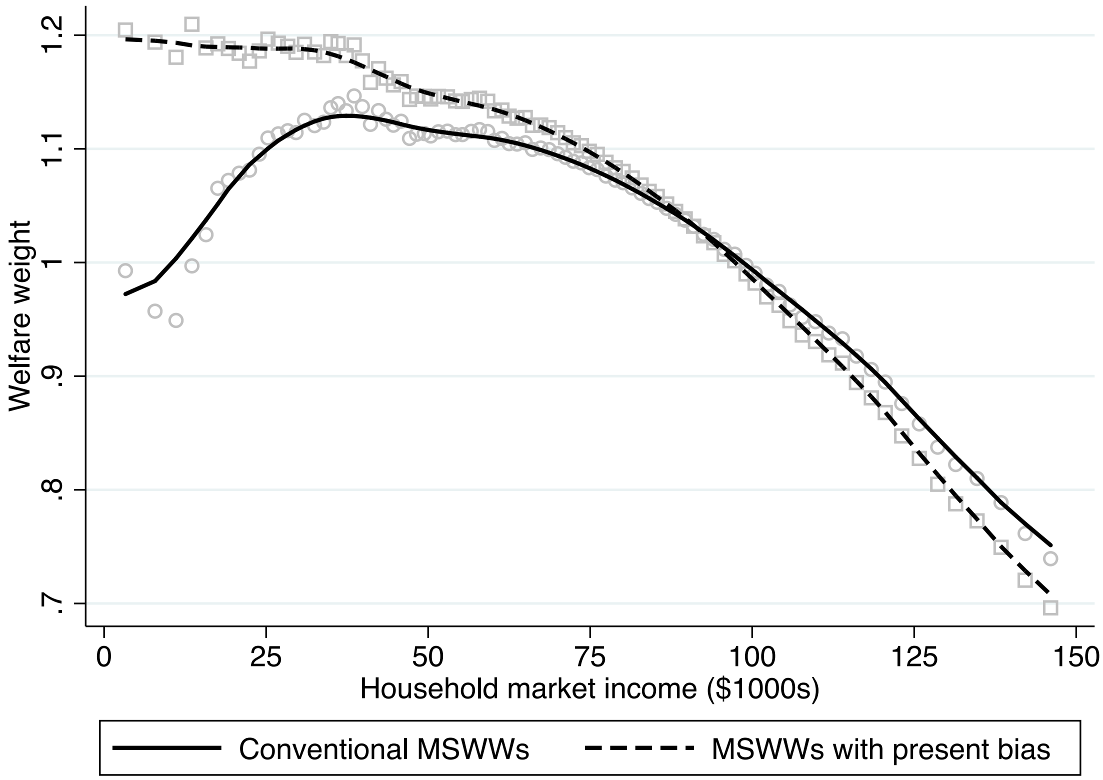

Figure 8: Welfare weights implicit in US policy under the conventional assumption of perfect optimization, and under calibrated present bias. Weights are computed by smoothing the income density using a fifth-order polynomial regression, then computing weights locally within each percentile of the income distribution. Lines are generated using kernel regression with a bandwidth of $5000.

income distribution, while the line plots the smoothed relationship, using kernel regression with a bandwidth of $5000.

The schedule of weights computed without present bias are strikingly unconventional in two respects. First, they rise over a substantial range of low incomes, peaking at about $30,000 in annual earnings. Since these EITC recipients are primarily single mothers, this result indicates that according to the conventional model, policy is designed to benefit single mothers at the low middle range of the income distribution much more than very poor working mothers—a result at odds with conventional normative assumptions. Similarly, the demographic homogeneity within this group of EITC recipients suggests that multidimensional heterogeneity, as in Choné and Laroque (2010), are an unlikely explanation for the pattern.

The second unconventional feature of the welfare weights computed under the conventional model is their low *level* on the poorest EITC recipients. Indeed, the weights suggest that policy has a strong preference for allocating marginal consumption to median earners (with welfare weights above 1.1) than to those with the very lowest incomes.

As shown by the dashed line in Figure 8, these unconventional features disappear when a calibrated degree of present bias is incorporated into the calculation of welfare weights. Specifically, weights are substantially higher than 1 at the bottom of the distribution, and decline monotonically with income.[24]

These results capture the sense in which the EITC can be explained by modest (but conventionally shaped) redistributive preferences. The dashed line in Figure 8 do not entail a strong desire to redistribute across low earners—indeed those with the lowest incomes have weights only slightly above the median—but the weights are nevertheless decreasing, consistent with the conventional objective of redistributing toward lower ability individuals.

7 Discussion

7.1 Implications for policy design

The analysis from the preceding sections allows the government only a single policy instrument: a nonlinear income tax. This restriction is useful for deriving policy implications that can realistically be implemented, avoiding complicated or unrealistic history-dependent policies that might arise from a full dynamic model. On the other hand, this constrained environment ignores some possible instruments which might be realistically feasible and useful for tailoring policies more specifically for present biased individuals. Here I speculate about two such instruments.

Timing of EITC payments. The current EITC is paid in aggregate at the end of the tax year—a structure that has generated mixed reactions. On one hand, spreading the credit across more frequent installments would help smooth consumption across the year and potentially alleviate liquidity constraints. Many EITC recipients carry substantial credit card balances, for example, and more frequent EITC payments would provide additional liquidity and could reduce costly

[24] A third possibly unconventional feature, common to both schedules, is the rather high welfare weight on high earners, even though many familiar utility functions have the marginal utility of consumption declining to zero as income grows large. The analysis of revealed redistributive preferences for high earners is not the focus of this paper, but see Lockwood and Weinzierl (2014) for a discussion.

interest expenses. On the other hand, the lump sum nature of the current EITC provides a short-run forced savings mechanism, and recipients often use the large annual payment to invest in durable goods—saving for which might otherwise prove difficult. Indeed, anecdotal evidence suggests EITC recipients do not *want* to receive distributed payments (Halpern-Meekin et al., 2015), a finding consistent with the very low uptake of the "Advance EITC" option, which allowed for more frequent payments. (See Romich and Weisner (2000) for a discussion, and Jones (2010) for experimental evidence of low desire for the Advance EITC.)

The model in Section 2.3 suggests that it is beneficial to levy taxes with a lag—thus the delayed nature of the present EITC is not as detrimental for combatting present bias as initial intuition might suggest. However that model recommends a rather different structure than the current, very lumpy schedule of EITC payments. Rather than paying subsidies at the end of each year, subsidies should be smoothed with a small, constant delay—perhaps on the order of one month. (Indeed, Kaur et al. (2015) suggests substantial discounting occurs within one week.)

How can this policy implication be squared with the seemingly desirable forced savings function of the current EITC? First note that existence of present bias is also consistent with the low observed uptake of the Advance EITC, and the lack of interest in up-front payments. As is well appreciated, theory predicts that sophisticated present biased agents will demand forced savings mechanisms. By paying the credit as an annual lump sum, the current EITC provides such a mechanism.[25] Nevertheless, whatever the merits of forced savings vehicles, there is no obvious benefit from bundling it with the corrective subsidy of the EITC. An unbundled policy would still allow sophisticated individuals to benefit from the EITC's corrective subsidy while choosing to divert a flexible share of income to a forced savings vehicle. On the other hand, naive individuals, who would not use the forced savings vehicle (and who might by repelled by that aspect of the current EITC) would, under an unbundled policy, nevertheless benefit from the corrective work subsidy.

This result sheds light on a number of policy discussions which have proposed reforming the EITC to provide more frequent payments. The Center for Economic Progress is currently exploring

[25] Many EITC recipients make use of refund anticipation loans, at high implicit interest rates, to avoid the typical three to six week processing delay. If the "long run self" chooses the EITC Advance enrollment, while the "short run self" decides whether to take out such a loan, both low EITC Advance uptake and widespread loan use are consistent with a model of sophisticated present bias.

such a reform in its Chicago Periodic EITC Payment Pilot, wherein EITC payments are distributed quarterly rather than annually[26] and presidential candidate Marco Rubio has proposed replacing the EITC with a wage enhancement program which would effectively include marginal work subsidies directly in recipients' paychecks.[27] The logic in this paper suggests the former approach is likely to provide a more targeted and efficient correction of present bias, while also providing EITC-receiving households with greater liquidity throughout the year than the present annual EITC.

Benefits waiting periods. A second possible dimension of fine tuning policy involves the timing of "lump sum" like benefits, $-T(0)$ in the notation of the models from Sections 2 and 3. In practice, these payments typically vary across programs. One feature that may reduce their distortions, particularly in contexts where workers may be tempted to quit and immediately draw benefits, is a required "waiting period" before benefits can be claimed. Under the model with commitment from Section 3, for example, a larger grant $-T(0)$ undermines productive commitments between firms and workers—however if the grant were delayed, so that self 1 discounts it by β, then it would appear a less tempting alternative. (Of course, such a waiting period must be weighed against the harm of delaying benefits for recipients.)

This may provide an economic rationale for the substantial delays associated with qualifying for programs such as disability insurance. Although this paper abstracted from many complexities of disability insurance by assuming that disabled workers could be screened perfectly, if screening is imperfect, then such delays may help prevent present biased workers from enrolling inefficiently.

7.2 Limitations

Although this paper attempts to present a believable calibration of the economy, it abstracts from several important complexities. Here I discuss several simplifying assumptions and their implications.

No human capital acquisition. Although I allow for delayed benefits from labor effort, I assume these benefits are separable from later labor supply decisions. This effectively rules out considerations of human capital accumulation. I make this assumption for a number of reasons.

[26]See `http://www.economicprogress.org/content/rethinking-eitc` for details.

[27]See `http://taxfoundation.org/blog/marco-rubio-proposes-replacement-earned-income-tax-credit`.

The first is pragmatic: the complexities of adequately accounting for human capital acquisition in optimal taxation are substantial, even without incorporating misoptimization (Stantcheva, 2014). For the sake of simplicity and transparency, and to generate results comparable to the existing literature exploring the (sub)optimality of negative marginal tax rates, I restrict consideration to a static distribution of ability. Second, the model in this paper still allows for a reduced-form relationship between human capital and present bias by calibrating bias conditional on income. Indeed, results in Section 4 suggest that bias is concentrated among those with lowest ability, as would be expected if some component of ability variation is due to human capital in which present biased individuals underinvest. Third, there is some empirical evidence that individuals who randomly receive work subsidies do not experience *persistent* increases in income relative to those who do not (Card and Hyslop, 2005, 2009)—a finding inconsistent with the notion that such subsidies raise human capital via on-the-job training effects.

To the extent that human capital effects are important for the design for optimal work subsidies, the directional implications are ambiguous. On one hand, some human capital is surely acquired on the job, raising the delayed benefits of work and likely inflating the size of optimal subsidies. On the other hand, human capital acquisition might other at times be a *substitute* for work with even more delayed benefits—for example, one may need to forego work to attend college. In that case work subsidies could exacerbate bias by discouraging human capital acquisition (see Shah and Steinberg (2015) for an application). This latter possibility may generate a rationale for exempting college-age individuals from EITC eligibility.

Perfectly competitive labor markets. In keeping with much of the optimal taxation literature, I assume that workers are employed in a perfectly competitive labor market, and that labor demand is infinitely elastic. This assumption has been questioned by Rothstein (2010), who argues that the incidence of work subsidies falls partly on employers. Finitely elastic labor demand undermines the argument for an EITC relative to guaranteed minimum income with high marginal tax rates, since the latter regime tends to reduce labor supply, raising wages and total transfers from employers to employees. Also in this vein, Kroft, Kucko, Lehmann and Schmieder (2015) incorporate endogenous wages and unemployment (not all job seekers find jobs) using a sufficient statistics approach. Their empirical results favor a negative income tax (rather than an EITC with negative marginal tax rates at low incomes) in a discrete model in the style of Saez (2002). Considerations of inelastic labor demand are beyond the scope of this paper, though they represent a natural extension, particularly for exploring the implications for the firm side of the economy with endogenous commitment contracts.

Fixed present bias. I assume throughout the model and calibrations that although present bias may vary with ability, it is fixed within individuals. In practice, biases may be mutable. One possibility, for example, is that exposure to the costs of present bias might lead individuals to improve their self control—a channel which would undermine the optimality of corrective subsidies that dampen such exposure. Another possibility, explored by Mullainathan et al. (2012) and Mullainathan and Shafir (2013), is that the conditions of poverty exacerbate behavioral biases. This possibility could be accommodated by writing bias as a decreasing function of consumption—a modification which would favor raising degree of overall redistribution. As with the case of human capital accumulation, this model would predict that temporary work subsidies should have persistent impacts on labor supply, inconsistent with the findings by Card and Hyslop cited above. Additionally, recent work by Carvalho et al. (2014) suggests that although liquidity constraints exacerbate measured present bias over monetary payments, they do not affect present bias over labor effort—consistent with a stable degree of structural bias. Still, optimal taxation with endogenous present bias is a promising avenue for further exploration.

8 Conclusion

As the study of optimal taxation begins to account for imperfect rationality and behavioral biases, a critical challenge is to quantify misoptimization accurately. This paper attempts progress by focusing on a particularly robust and well understood source of misoptimization—present bias—which, recent evidence suggests, generates substantial labor supply distortions.

A theoretical model of optimal taxation with present bias generates new theoretical implications, including a "negative at the bottom" result and a surprising implication for optimal tax timing: if some individuals are paid up-front, then it is beneficial to levy taxes with a delay. A model with sophisticated agents and endogenous commitment contracts generates novel predictions for labor supply patterns—a U-shaped schedule of elasticities, and an inverse U-shaped schedule of income effects—which are consistent with patterns in the empirical literature.

A compilation of existing estimates of present bias provides strong evidence of such bias at a structural level. To allow for the possibility that bias is corrected by commitment contracts, I present new estimates of "residual bias" based on trends in subjective well-being responses following the 1990s welfare reforms in the US. Estimates of structural bias and residual bias are similar, and suggest bias is heavily concentrated at low incomes. Although estimates of misoptimization will surely continue to improve, the consistency of results across methodologies provides some hope that misoptimization can be estimated with sufficient precision to provide clear guidance for policy design.

The implications for optimal tax policy depend on one's view of optimal redistribution. If redistributive tastes are modest, like those in the baseline calibrations in this paper, then the negative marginal tax rates generated by the EITC may be close to optimal. In that case the policy implications of these results are clear: the existing EITC should not be reduced or greatly reformed—indeed, it might beneficially be extended to workers without children, and made more salient as a wage subsidy included directly with one's paycheck. On the other hand, if redistributive preferences are stronger, in line with the welfare weights often assumed by optimal tax theorists (such as logarithmic utility of consumption) then negative marginal tax rates at low incomes appear to be suboptimal even in the context of present biased workers. As a positive matter, the widely noted fact that tax rates on high earners appear to embody substantially weaker redistributive concerns than conventional utility functions provides some indication that present bias may be a reasonable explanation for the shape of the existing EITC.

References

Allcott, Hunt and Taubinsky, Dmitry. (forthcoming). 'Evaluating behaviorally-motivated policy: Experimental evidence from the lightbulb market', *American Economic Review* .

Augenblick, Ned, Niederle, Muriel and Sprenger, Charles. (2013), Working over time: Dynamic inconsistency in real effort tasks.

Beaudry, Paul, Blackorby, Charles and Szalay, Dezsö. (2009). 'Taxes and Employment Subsidies in Optimal Redistribution Programs', *The American Economic Review* 99(1), 216.

Bernheim, B Douglas and Rangel, Antonio. (2007). 'Toward choice-theoretic foundations for behavioral welfare economics', *The American economic review* pp. 464–470.

Bernheim, B Douglas and Rangel, Antonio. (2009). 'Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics', *The Quarterly Journal of Economics* 124(1), 51–104.

Besley, Timothy and Coate, Stephen. (1992). 'Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs', *American Economic Review* 82(1), 249–261.

Besley, Timothy and Coate, Stephen. (1995). 'The design of income maintenance programmes', *The Review of Economic Studies* 62(2), 187–221.

Blundell, Richard and Shephard, Andrew. (2011). 'Employment, hours of work and the optimal taxation of low-income families', *The Review of Economic Studies* p. rdr034.

Bourguignon, François and Spadaro, Amedeo. (2012). 'Tax–benefit revealed social preferences', *The Journal of Economic Inequality* 10(1), 75–108.

Boyd-Swan, Casey, Herbst, Chris M, Ifcher, John and Zarghamee, Homa. (2013). 'The earned income tax credit, health, and happiness'.

Card, David and Hyslop, Dean R. (2005). 'Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers', *Econometrica* 73(6), 1723–1770.

Card, David and Hyslop, Dean R. (2009). 'The dynamic effects of an earnings subsidy for long-term welfare recipients: Evidence from the self sufficiency project applicant experiment', *Journal of Econometrics* 153(1), 1–20.

Carvalho, Leandro S, Meier, Stephan and Wang, Stephanie W. (2014). 'Poverty and economic decision making: evidence from changes in financial resources at payday', *Unpublished manuscript* .

Center on Budget and Policy Priorities. (2014), Policymakers Often Overstate Marginal Tax Rates for Lower-Income Workers and Gloss Over Tough Trade-Offs in Reducing Them, Technical report, http://www.cbpp.org/research/policymakers-often-overstate-marginal-tax-rates-for-lower-income-workers-and-gloss-over.

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

Chetty, Raj. (2015). 'Behavioral Economics and Public Policy: A Pragmatic Perspective', *American Economic Review* 105(5), 1–33.

Chetty, Raj, Friedman, John N and Saez, Emmanuel. (2013). 'Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings', *American Economic Review* 103(7), 2683–2721.

Chetty, Raj, Looney, Adam and Kroft, Kory. (2009). 'Salience and Taxation: Theory and Evidence', *American Economic Review* 99(4), 1145–1177.

Choné, Philippe and Laroque, Guy. (2010). 'Negative marginal tax rates and heterogeneity', *The American Economic Review* pp. 2532–2547.

Congressional Budget Office. (2012), Effective Marginal Tax Rates for Low- and Moderate-Income Workers, Technical report, https://www.cbo.gov/publication/43709.

Cuff, Katherine. (2000). 'Optimality of workfare with heterogeneous preferences', *Canadian Journal of Economics/Revue canadienne d'économique* 33(1), 149–174.

Dahl, Gordon B and Lochner, Lance. (2012). 'The impact of family income on child achievement: Evidence from the earned income tax credit', *The American Economic Review* 102(5), 1927–1956.

DellaVigna, Stefano. (2009). 'Psychology and Economics: Evidence from the Field', *Journal of Economic Literature* 47(2), 315–372.

DellaVigna, Stefano, Lindner, Attila, Reizer, Balázs and Schmieder, Johannes F. (2015). 'Reference-Dependent Job Search: Evidence from Hungary', *Unpublished Working Paper* .

DellaVigna, Stefano and Paserman, M Daniele. (2005). 'Job Search and Impatience', *Journal of Labor Economics* 23(3).

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

Diamond, Peter A. (1998). 'Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates', *American Economic Review* 88(1), 83–95.

Di Tella, Rafael, MacCulloch, Robert J and Oswald, Andrew J. (2001). 'Preferences over inflation and unemployment: Evidence from surveys of happiness', *American Economic Review* 91(1), 335–341.

Drenik, Andres and Perez-Truglia, Ricardo. (2014). 'Sympathy for the Diligent and the Demand for Workfare', *Available at SSRN 1599368* .

Fang, Hanming and Silverman, Dan. (2009). 'Time-inconsistency and welfare program participation: Evidence from the nlsy*', *International Economic Review* 50(4), 1043–1077.

Farhi, Emmanuel and Gabaix, Xavier. (2015), Optimal Taxation with Behavioral Agents.

Fleurbaey, Marc and Maniquet, Francois. (2006). 'Fair income tax', *The Review of Economic Studies* 73(1), 55–83.

Gerritsen, Aart. (2015). 'Optimal taxation when people do not maximize well-being'.

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

Halpern-Meekin, Sarah, Edin, Kathryn, Tach, Laura and Sykes, Jennifer. (2015), *It's Not Like I'm Poor*, University of California Press.

Hellwig, Martin F. (2007). 'A contribution to the theory of optimal utilitarian income taxation', *Journal of Public Economics* 91(7), 1449–1477.

Hendren, Nathaniel. (2014). 'The Inequality Deflator: Interpersonal Comparisons without a Social Welfare Function'.

Herbst, Chris M. (2013). 'Welfare reform and the subjective well-being of single mothers', *Journal of Population Economics* 26(1), 203–238.

Ifcher, John. (2010). 'The happiness of single mothers after welfare reform', *The BE Journal of Economic Analysis & Policy* 11(1).

Ifcher, John and Zarghamee, Homa. (2014). 'The Happiness of Single Mothers: Evidence from the General Social Survey', *Journal of Happiness Studies* 15(5), 1219–1238.

Jacquet, Laurence and Lehmann, Etienne. (2014), Optimal Nonlinear Income Taxation with Multidimensional Types: The Case with Heterogeneous Behavioral Responses.

Jacquet, Laurence, Lehmann, Etienne and Van der Linden, Bruno. (2013). 'Optimal redistributive taxation with both extensive and intensive responses', *Journal of Economic Theory* 148(5), 1770–1805.

Jones, Damon. (2010). 'Information, Preferences, and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401 (k) Savings', *American Economic Journal: Applied Economics* pp. 147–163.

Jones, Damon and Mahajan, Aprajit. (2015). 'Time-Inconsistency and Savings: Experimental Evidence from Low-Income Tax Filers', *Center for Financial Security Working Paper* .

Kanbur, Ravi, Keen, Michael and Tuomala, Matti. (1994). 'Optimal non-linear income taxation for the alleviation of income-poverty', *European Economic Review* 38(8), 1613–1632.

Kaplow, Louis. (2007). 'Optimal income transfers', *International Tax and Public Finance* 14(3), 295–325.

Kaur, Supreet, Kremer, Michael and Mullainathan, Sendhil. (2015). 'Self-control at work', *Forthcoming* .

Kotlikoff, Laurence and Rapson, David. (2009). 'Does it pay, at the margin, to work and save?', *Tax Policy and the Economy* 21.

Kroft, Kory, Kucko, Kavan J, Lehmann, Etienne and Schmieder, Johannes F. (2015), Optimal Income Taxation with Unemployment and Wage Responses: A Sufficient Statistics Approach, Technical report, National Bureau of Economic Research.

Laibson, David. (1997). 'Golden eggs and hyperbolic discounting', *The Quarterly Journal of Economics* 112(2), 443–478.

Laibson, David, Repetto, Andrea and Tobacman, Jeremy. (2007), Estimating discount functions with consumption choices over the lifecycle, Technical report, National Bureau of Economic Research.

Lockwood, Benjamin B and Taubinsky, Dmitry. (2015), Regressive Sin Taxes.

Lockwood, Benjamin B and Weinzierl, Matthew. (2014), Positive and Normative Judgments Implicit in U.S. Tax Policy, and the Costs of Unequal Growth and Recessions.

Lockwood, Benjamin B. and Weinzierl, Matthew. (2015). 'De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution', *Journal of Public Economics* 124(0), 74–80.
URL: *http://www.sciencedirect.com/science/article/pii/S0047272715000134*

McGinn, Kathleen L, Ruiz Castro, Mayra and Lingo, Elizabeth Long. (2015). 'Mums the Word! Cross-national Effects of Maternal Employment on Gender Inequalities at Work and at Home'.

Meier, Stephan and Sprenger, Charles. (2015). 'Stability of time preferences', *Review of Economics and Statistics* .

Meyer, Bruce D. (2010), The effects of the Earned Income Tax Credit and recent reforms, *in* 'Tax Policy and the Economy, Volume 24', The University of Chicago Press, pp. 153–180.

Meyer, Bruce D and Sullivan, James X. (2008). 'Changes in the consumption, income, and well-being of single mother headed families', *The American Economic Review* pp. 2221–2241.

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

Moffitt, Robert. (2006). 'Welfare work Requirements with Paternalistic Government Preferences*', *The Economic Journal* 116(515), F441–F458.

Mullainathan, Sendhil, Schwartzstein, Joshua and Congdon, William J. (2012). 'A reduced-form approach to behavioral public finance', *Economics* 4.

Mullainathan, Sendhil and Shafir, Eldar. (2013), *Scarcity: Why having too little means so much*, Macmillan.

Paserman, M Daniele. (2008). 'Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation*', *The Economic Journal* 118(531), 1418–1452.

Romich, Jennifer L and Weisner, Thomas. (2000). 'How families view and use the EITC: Advance payment versus lump sum delivery', *National tax journal* pp. 1245–1265.

Rothstein, Jesse. (2010). 'Is the EITC as good as an NIT? Conditional cash transfers and tax incidence', *American Economic Journal: Economic Policy* pp. 177–208.

Saez, Emmanuel. (2001). 'Using elasticities to derive optimal income tax rates', *The Review of Economic Studies* 68(1), 205–229.

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

Saez, Emmanuel, Slemrod, Joel and Giertz, Seth H. (2012). 'The elasticity of taxable income with respect to marginal tax rates: A critical review', *Journal of Economic Literature* 50(1), 3–50.

Seade, Jesus. (1982). 'On the sign of the optimum marginal income tax', *The Review of Economic Studies* 49(4), 637–643.

Seade, Jesus K. (1977). 'On the shape of optimal tax schedules', *Journal of public Economics* 7(2), 203–235.

Shah, Anuj K, Mullainathan, Sendhil and Shafir, Eldar. (2012). 'Some consequences of having too little', *Science* 338(6107), 682–685.

Shah, Manisha and Steinberg, Bryce Millett. (2015), Workfare and Human Capital Investment: Evidence from India.

Spinnewijn, Johannes. (2014). 'Unemployed but optimistic: Optimal insurance design with biased beliefs', *Journal of the European Economic Association* .

Stantcheva, Stefanie. (2014), Optimal Taxation and Human Capital Policies over the Life Cycle.

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

Vickrey, William. (1945). 'Measuring marginal utility by reactions to risk', *Econometrica: Journal of the Econometric Society* pp. 319–333.

Werning, Ivan. (2000), An Elementary Proof of Positive Optimal Marginal Taxes.

Online appendix

Appendix A Proof of Proposition 1 using the Hamiltonian

Proof. Let $\ell(w) = z(w)/w$, and let $V(w)$ denote rescaled decision utility for ability w:

$$V(w) := w\ell(w) - T(w(\ell(w)) - \frac{v(\ell(w))}{\beta(w)}.$$

The individual's optimization implies

$$w(1 - T'(w\ell(w))) = \frac{v'(\ell(w))}{\beta(w)},$$

and so we have

$$V'(w) = \frac{\ell(w)v'(\ell(w))}{\beta(w)w} + \frac{v(\ell(w))\beta'(w)}{\beta(w)^2}.$$

Then experienced utility is equal to

$$V(w) + \left(\frac{1-\beta(w)}{\beta(w)}\right) v(\ell(w)),$$

and, in a modified version of the standard optimal control setup, we can take $V(w)$ as the state variable $\ell(w)$ as the control variable, and write the problem as

$$\max \int_{w_0}^{w_1} \alpha(w) \left(V(w) + \left(\frac{1-\beta(w)}{\beta(w)}\right) v(\ell(w)) \right) f(w)dw$$

subject to

$$\int_{w_0}^{w_1} \left(V(w) + \frac{v(\ell(w))}{\beta(w)} \right) f(w)dw \leq \int_{w_0}^{w_1} w\ell(w)f(w)dw - E. \tag{17}$$

and

$$V'(w) = \frac{\ell(w)v'(\ell(w))}{\beta(w)w} + \frac{v(\ell(w))\beta'(w)}{\beta(w)^2}. \tag{18}$$

Letting λ denote the multiplier on the budget constraint in (17) $m(w)$ denote the multipliers on the constraint in (18), then the Hamiltonian for this problem is

$$\mathcal{H} = \left[\alpha(w) \left(V(w) + \left(\frac{1-\beta(w)}{\beta(w)}\right) v(\ell(w)) \right) - \lambda \left(V(w) + \frac{v(\ell(w))}{\beta(w)} - w\ell(w) \right) \right] f(w) + m(w) \left(\frac{\ell(w)v'(\ell(w))}{\beta(w)w} + \frac{v(\ell(w))\beta'(w)}{\beta(w)^2} \right).$$

The usual solution technique requires

$$m'(w) = -\frac{\partial \mathcal{H}}{\partial V} = (\lambda - \alpha(w))\, f(w).$$

Maximizing $\mathcal{H}$ with respect to $\ell(w)$, we have

$$\left(-\alpha(w)\left(\frac{1-\beta(w)}{\beta(w)}\right)v'(\ell(w))+\lambda\left(\frac{v'(\ell(w))}{\beta(w)}-w\right)\right)f(w)$$
$$=m(w)\left(\frac{v'(\ell(w))+\ell(w)v''(\ell(w))}{\beta(w)w}+\frac{v'(\ell(w))\beta'(w)}{\beta(w)^2}\right). \quad (19)$$

Using the fact that $m(w_1)=0$ (no distortion at the top) we have

$$m(w)=\int_{w_1}^{w}m'(w)dw=-\int_{w}^{w_1}m'(w)dw=\int_{w}^{w_1}(\alpha(w)-\lambda)\,f(w)dw.$$

Substituting into (19) and rearranging,

$$\frac{T'}{1-T'}=$$
$$\frac{1}{f(w)}\int_{w}^{w_1}(1-g(w))\,f(w)dw\left(\frac{1+\ell(w)v''(\ell(w))/v'(\ell(w))}{w}+\frac{\beta'(w)}{\beta(w)}\right)-g(w)\,(1-\beta(w))\,.$$

Then substituting in the expressions (from the text) for the elasticities of labor supply and present bias yields the expression in Proposition 1. □

Appendix B Elasticity concepts

In this section I define formally the elasticities employed in Section 2.2. As in Jacquet and Lehmann (2014), these include "circularities" due to the curvature of the tax function, wherein an initial change causes a change in earnings, leading to a further change in the tax rate, etc. To this, it is useful to parameterize a local tax reform about earnings level z^* using r and τ, where r is a reduction in the level of taxes at z^* (identifying an income effect) and τ is a compensated change in the marginal tax rate at z^* (identifying a compensated elasticity):

$$\hat{T}(z;z^*,r,\tau):=T(z)-r+\tau(z-z^*).$$

Then let $\hat{z}(i;r,\tau)$ denote i's optimal choice of earnings as a function of the reform parameters r and τ, when the tax code is perturbed around i's status quo chosen level of earnings $z(i)$:

$$\hat{z}(i;r,\tau):=\arg\max_{z}\left\{-v(z/w_i)-\chi_i\cdot 1\{z>0\}+\beta_i u(z-\hat{T}(z;z(i),r,\tau))\right\}.$$

The compensated elasticity is defined as the elasticity of income with respect to the marginal tax rate increase, τ, evaluated at the status quo:

$$\varepsilon(i):=\left(-\frac{d\hat{z}(i;0,\tau)}{d\tau}\bigg|_{\tau=0}\right)\frac{1-T'(z(i))}{z(i)}.$$

Computing this derivation using the separable form in Equation (1), we find that for any labor force participant,

$$\varepsilon(i)=\frac{1/z(i)}{\frac{v''(z(i)/w_i)/w_i}{v'(z(i)/w_i)}+\frac{T''(z(i))}{1-T'(z(i))}-\frac{u''(z(i)-T(z(i)))}{u'(z(i)-T(z(i)))}(1-T'(z(i))}.$$

An implication of this result is that if labor supply disutility is isoelastic (i.e., $v(x) = ax^b$ for constants a and b) then the elasticity is constant conditional on income, as $\frac{v''(z(i)/w_i)/w_i}{v'(z(i)/w_i)} = \frac{b-1}{z(i)}$. That is, different combinations of ability and bias which give rise to the same chosen level of earnings also give rise to the same compensated elasticity of earnings.

The income effect is defined formally as

$$\eta(i) := \left(-\frac{d\hat{z}(i;r,0)}{dr}\bigg|_{r=0} \right) \left(1 - T'(z(i))\right).$$

Again computing the derivative using Equation (1), we have

$$\eta(i) = \frac{-u'((z(i) - T(z(i)))^{-1}}{\frac{v''(z(i)/w_i)/w_i}{v'(z(i)/w_i)} + \frac{T''(z(i))}{1-T'(z(i))} - \frac{u''(z(i)-T(z(i)))}{u'(z(i)-T(z(i)))}(1 - T'(z(i))}(1 - T'(z(i))).$$

As with the compensated elasticity, this implies that if disutility of labor is isoelastic, then the income effect is the same for all individuals with a given level of earnings.

Since the participation elasticity is discrete, it is not defined as a function of individual type i, but rather as a function of income, representing the number of individuals who enter the labor force to earn z when the level of taxes at z is reduced slightly. Let $z^p(i)$ denote the optimal *participation earnings* of i—the earnings chosen if $z = 0$ is were removed from the budget set—under T, and $\hat{z}^p(i;r,\tau)$ the participation earnings under the reformed $\hat{T}$.[28] The participation elasticity quantifies the change in the share of labor force participants with a given participation earnings z^* in response to a small tax reduction at z^*. To write this formally, note that by construction $h(z^*)$ is equal to the measure of individuals with participation earnings $z^p(i) = z^*$ who actually participate:

$$h(z^*) = \int_{\mathcal{I}^p(z^*)} 1\left\{\chi_i < -v\left(\frac{z^p(i)}{w_i}\right) + \beta_i \left(u\left(z^p(i) - T(z^p(i))\right) - u(-T(0))\right)\right\} d\mu(i)$$

Now define $\hat{h}(z^*;r,\tau)$ to be the measure of individuals with participation earnings $z^p(i) = z^*$ under the status quo tax $T(z)$ who participate under the reformed tax $\hat{T}(z;z^*,r,\tau)$:

$$\hat{h}(z^*;r,\tau) =$$
$$\int_{\mathcal{I}^p(z^*)} 1\left\{\chi_i < -v\left(\frac{\hat{z}^p(i;r,\tau)}{w_i}\right) + \beta_i \left(u\left(\hat{z}^p(i;r,\tau) - \hat{T}(\hat{z}^p(i;r,\tau);z^*,r,\tau)\right) - u(-T(0))\right)\right\} d\mu(i).$$

Note that the set $I^p(z^*)$ is defined with respect to the status quo tax $T(z)$ and does not vary with the tax reform—this is because we are interested in the change in the share of labor force participants due solely to entry from outside the labor force, and not from the change in individuals earning z^* due to local adjustments in response to the reform.

Using this notation, the participation elasticity at z^* is defined as

$$\rho(z) := \left(\frac{d\hat{h}(z;r,0)}{dr}\bigg|_{r=0} \right) \frac{z - T(z) + T(0)}{h(z)}.$$

[28] That is, $\hat{z}^p(i;r,\tau) := \arg\max_z \left\{-v(z/w_i) + \beta_i u(z - \hat{T}(z;z(i),r,\tau))\right\}$.

Appendix C Derivation of Welfare Internality Terms

The change in welfare due to the welfare internality from the substitution effect (the local adjustment of earnings among individuals who face a higher marginal tax rate) is equal to

$$dS_W = d\tau\epsilon\left(\frac{1}{\lambda}\right)\int_{\mathcal{I}(z^*)}\alpha_i\left(\left.\frac{d\hat{z}(i;0,\tau)}{d\tau}\right|_{\tau=0}\right)\left[-v'(z^*/w_i)/w_i + u'(z^*-T(z^*))(1-T'(z^*))\right]d\mu(i).$$

This is the change in experienced utility among individuals earning z^* whose earnings respond locally due to the rise in the marginal tax rate at z^*. The division by λ converts from utility (measured in the integral) into dollars, for summation with revenue effects. Individual optimization implies that $-v'(z^*/w_i)/w_i + \beta_i u'(z^*-T(z^*))(1-T'(z^*)) = 0$; subtracting this from the term in the integral yields

$$\begin{aligned}
dS_W &= d\tau\epsilon\int_{\mathcal{I}(z^*)}\left(\left.\frac{d\hat{z}(i;0,\tau)}{d\tau}\right|_{\tau=0}\right)(1-\beta_i)\left(\frac{\alpha_i u'(z^*-T(z^*))}{\lambda}\right)(1-T'(z^*))d\mu(i)\\
&= d\tau\epsilon\int_{\mathcal{I}(z^*)}\left(\left.\frac{d\hat{z}(i;0,\tau)}{d\tau}\right|_{\tau=0}\right)(1-\beta_i)\,g(i)(1-T'(z^*))d\mu(i)\\
&= d\tau\epsilon z^*\int_{\mathcal{I}(z^*)}-\varepsilon(i)\,(1-\beta_i)\,g(i)d\mu(i).
\end{aligned}$$

Rewriting the integral as an expectation yields the expression in the text. The derivation of dI_W is analogous.

The change in welfare due to the welfare internality on the extensive margin from a reduction in taxes dr for an income band of width $\epsilon \to 0$ around income level z^* is

$$dr\epsilon\left(\frac{1}{\lambda}\right)\int_{\mathcal{I}_{ext}(z^*)}\alpha_i\left[-v(z^*/w_i)-\chi_i+u(z^*-T(z^*))-u(-T(0))\right]d\mu(i).$$

By individual optimization, individuals on the extensive margin (i.e., those in $\mathcal{I}_{ext}(z^*)$) satisfy $-v(z^*/w_i)-\chi_i+\beta_i u(z^*-T(z^*))-\beta_i u(-T(0)) = 0$. Subtracting this from the term in brackets yields

$$dr\epsilon\left(\frac{1}{\lambda}\right)\int_{\mathcal{I}_{ext}(z^*)}\alpha_i(1-\beta_i)\left[u(z^*-T(z^*))-u(-T(0))\right]d\mu(i) =$$
$$dr\epsilon\left[z^*-T(z^*)+T(0)\right]\int_{\mathcal{I}_{ext}(z^*)}g_{ext}(i)(1-\beta_i)d\mu(i),$$

employing the definition of the extensive margin welfare weight from the text. Rewriting the integral as a conditional expectation, and using the fact that $\int_{\mathcal{I}_{ext}(z^*)}d\mu(i) = \left.\frac{d\hat{h}(z;r,0)}{dr}\right|_{r=0}$, this expression can be written

$$dr\epsilon\rho(z^*)h(z^*)\mathbb{E}\left[g_{ext}(i)(1-\beta_i)|i\in\mathcal{I}_{ext}(z^*)\right].$$

Under the perturbation in the text, this extensive margin internality effect occurs at all incomes above the point of the marginal tax reform. Integrating across those incomes yields the expression in the text.

Appendix D Proof of Lemma 1

Proof. Suppose that for some χ' and χ'', $Y(\chi') = Y(\chi'') = 0$. Then by (10), $\beta\tilde{u}(Z(\chi') = \beta\tilde{u}(Z(\chi''))$ (using the inequality in both directions), and therefore $Z(\chi') = Z(\chi'')$. Thus any individuals who do not work must receive the same compensation—call this value Z_0. Now suppose there is some type χ^* such that $Y(\chi^*) > 0$, and consider two types $\chi^\dagger$, $\chi^{\dagger\dagger}$, both less than χ^*. If any types do not work, then (10) requires $-v(Y(\chi^*)) - \chi^* + \beta\tilde{u}(Z(\chi^*)) \geq -v(0) + \beta\tilde{u}(Z_0)$. Since $\chi^\dagger < \chi^*$, it follows that $\chi^\dagger$ would rather mimic χ^* than mimic a non-worker, implying that $Y(\chi^\dagger) > 0$, and similarly $Y(\chi^{\dagger\dagger}) > 0$. Thus we have that if some χ^* works, all types $\chi < \chi^*$ also work. Let $\mu(\chi|\chi')$ denote the multiplier on the constraint (10), requiring that type χ' not prefer to mimic χ, and let λ denote the multiplier on (11). Then the derivative of the Lagrangian from (9)–(12) with respect to $Y(\chi^\dagger)$ is

$$\frac{\partial \mathcal{L}}{\partial Y(\chi^\dagger)} = -\tilde{v}'(Y(\chi^\dagger) + \lambda + \int \mu(\chi^\dagger|\chi)\tilde{v}'(Y(\chi^\dagger))d\chi - \int \mu(\chi|\chi^\dagger)\tilde{v}'(Y(\chi^\dagger))d\chi = 0, \tag{20}$$

and

$$\frac{\partial \mathcal{L}}{\partial Y(\chi^{\dagger\dagger})} = -\tilde{v}'(Y(\chi^{\dagger\dagger}) + \lambda + \int \mu(\chi^{\dagger\dagger}|\chi)\tilde{v}'(Y(\chi^{\dagger\dagger}))d\chi - \int \mu(\chi|\chi^{\dagger\dagger})\tilde{v}'(Y(\chi^{\dagger\dagger}))d\chi = 0. \tag{21}$$

By (10), $-v(Y(\chi^\dagger) + \beta\tilde{u}(Z(\chi^\dagger)) = -v(Y(\chi^{\dagger\dagger}) + \beta\tilde{u}(Z(\chi^{\dagger\dagger}))$. Therefore $\mu(\chi^\dagger|\chi') = \mu(\chi^{\dagger\dagger}|\chi')$ for all χ' (it is equally tempting for other types to mimic $\chi^\dagger$ and $\chi^{\dagger\dagger}$). Moreover, $\mu(\chi'|\chi^\dagger) = \mu(\chi'|\chi^{\dagger\dagger})$ for all χ' such that $Y(\chi') > 0$—it is equally tempting for $\chi^\dagger$ and $\chi^{\dagger\dagger}$ to mimic some other working types—and the assumption that $\chi^{\dagger\dagger\dagger}$ is greater than and $\chi^\dagger$ and $\chi^{\dagger\dagger}$ implies that $Y = 0$ is strictly more tempting for $\chi^{\dagger\dagger\dagger}$ than for $\chi^\dagger$ or $\chi^{\dagger\dagger}$. Therefore multipliers on the constraints preventing $\chi^\dagger$ and $\chi^{\dagger\dagger}$ from mimicking non-workers are all zero. Thus $\mu(\chi'|\chi^\dagger) = \mu(\chi'|\chi^{\dagger\dagger})$ for all χ'. Thus rearranging (20) and (21) yields

$$\tilde{v}'(Y(\chi^\dagger)) = \frac{\lambda}{1 - \int \mu(\chi^\dagger|\chi)d\chi + \int \mu(\chi|\chi^\dagger)d\chi} = \frac{\lambda}{1 - \int \mu(\chi^{\dagger\dagger}|\chi)d\chi + \int \mu(\chi|\chi^{\dagger\dagger})d\chi} = \tilde{v}'(Y(\chi^{\dagger\dagger})).$$

Thus if some χ^* works, then all types $\chi < \chi^*$ also work and have the same production and compensation bundle denoted Y^* and Z^*. Moreover, if some χ^{**} does not work, then all types $\chi > \chi^{**}$ do not work, so $Y(\chi) = 0$, and all receive the same compensation Z_0. Since the support of χ is connected, this implies there is some threshold χ^* below which all types work, and above which all types do not. It is easily verified that χ^* must therefore be indifferent between working and not, implying $-v(Y^*) - \chi^* + \beta\tilde{u}(Z^*) = \beta\tilde{u}(Z_0)$. Thus the optimal contract is characterized by the parameters Z_0, Z^*, Y^*, and χ^*, and the lemma is proved. □

Appendix E Proof of Proposition 5

Proof. **Full commitment.** Let w^h solve (15) with equality when $Z_0 = z_0$:

$$-v(Y^*/w^h)/w^h - \overline{\chi} + \beta\tilde{u}(Z^*) = \beta\tilde{u}(z_0).$$

Then by letting $Y^* = Z^* - \kappa$, with $v'(Y^*/w^h)/w^h = \tilde{u}(Z^*)$, and $\chi^* = \overline{\chi}$, (13) is maximized subject to (14), and constraints (15) and (16) do not bind, demonstrating that this constitutes a constrained optimum. Since $\frac{v'(Y^*/w^h)/w^h}{u'(Z^* - T(Z^*))(1-T'(Z^*))} = \frac{v'(Y^*/w^h)/w^h}{\tilde{u}'(Z^*)} = 1$, the misoptimization wedge is zero. The same trivially holds for all $w > w^h$.

No commitment. Substituting (16) into (15) yields a combined constraint

$$-\tilde{v}(Y^*) - \chi^* + \beta\tilde{u}(Z^*) \geq \beta\tilde{u}(z_0).$$

Moreover, using the fact that $Z^* \leq Y^*$, we can write

$$-v(Y^*/w) - \chi^* + \beta\tilde{u}(Y^*) \geq -\tilde{v}(Y^*) - \chi^* + \beta\tilde{u}(Z^*) \geq \beta\tilde{u}(z_0).$$

Letting w grow arbitrarily small, either $v(Y^*/w)$ grows arbitrarily large, or Y^* becomes smaller than z_0. In either case, the inequality above is violated, showing that for sufficiently low w, no contract with an interior χ^* satisfies (14)–(16). Therefore the only (possibly) feasible contract is $\chi^* = \underline{\chi}$ and $Z_0 = -\kappa$. If this contract violates (16), it is infeasible; if not, it is nevertheless dominated by the the outside option of flexibly setting labor supply in period 1, which provides a lower bound of utility equal to $Z_0 = 0$.

Limited commitment. Let λ denote the multiplier on (14), and since (15) binds at the optimum if $\chi^* < \overline{\chi}$ and the constraint is slack if $\chi^* = \overline{\chi}$, we can suppose that Z_0 satisfies (15) with equality in all cases, and we can use that equation to implicitly determine χ^* as a function of Y^*, Z^*, and Z_0. Let μ denote the multiplier on (16). Then the optimal contract satisfies

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial Y^*} &= F(\chi^*)\left(-\tilde{v}'(Y^*) + \lambda\right) + \frac{\partial \mathcal{L}}{\partial \chi^*}\frac{\partial \chi^*}{\partial Y^*} \\ &= F(\chi^*)\left(-\tilde{v}'(Y^*) + \lambda\right) - f(\chi^*)\left((1-\beta)\left(\tilde{u}(Z^*) - \tilde{u}(Z_0)\right) + \lambda\left(Y^* - Z^* + Z_0\right)\right)\tilde{v}'(Y^*),\end{aligned}$$

implying

$$\tilde{v}'(Y^*) = \frac{\lambda}{1 + \frac{f(\chi^*)}{F(\chi^*)}\left((1-\beta)\left(\tilde{u}(Z^*) - \tilde{u}(Z_0)\right) + \lambda\left(Y^* - Z^* + Z_0\right)\right)}.$$

Similarly,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial Z^*} &= F(\chi^*)\left(\tilde{u}'(Z^*) - \lambda\right) + \frac{\partial \mathcal{L}}{\partial \chi^*}\frac{\partial \chi^*}{\partial Z^*} \\ &= F(\chi^*)\left(\tilde{u}'(Z^*) - \lambda\right) + f(\chi^*)\left((1-\beta)\left(\tilde{u}(Z^*) - \tilde{u}(Z_0)\right) + \lambda\left(Y^* - Z^* + Z_0\right)\right)\beta\tilde{u}'(Z^*),\end{aligned}$$

implying

$$\tilde{u}'(Z^*) = \frac{\lambda}{1 + \beta\frac{f(\chi^*)}{F(\chi^*)}\left((1-\beta)\left(\tilde{u}(Z^*) - \tilde{u}(Z_0)\right) + \lambda\left(Y^* - Z^* + Z_0\right)\right)}.$$

Thus at the constrained optimum,

$$\frac{\tilde{v}'(Y^*)}{\tilde{u}'(Z^*)} = \frac{1 + \beta A}{1 + A},$$

with $A = \frac{f(\chi^*)}{F(\chi^*)}\left((1-\beta)\left(\tilde{u}(Z^*) - \tilde{u}(Z_0)\right) + \lambda\left(Y^* - Z^* + Z_0\right)\right)$. Since $A > 0$, this ratio lies between 1 (when $A = 0$) and β (as $A \to \infty$), implying that the misoptimization wedge lies between 0 and $1 - \beta$. □

Appendix F Details of numerical example in Section 3

This example uses a quasilinear utility function:

$$U(c, y, w) = c - \frac{(y/w)^{(1+1/e)}}{1 + 1/e},$$

where e is the elasticity of labor supply. I set $e = 0.3$, and I assume a flat tax marginal tax rate of 0.3 with no lump sum grant. (The nature of the results below is not sensitive to these assumptions.) I assume a population with heterogenous ability, with w distributed uniformly on $[\underline{w}, \overline{w}]$, and I assume all individuals face the same distribution of fixed costs $F(\chi)$.

A simplifying assumption. The nature of the boundary between the "limited commitment" and "no commitment" contract regions depends on whether it is possible to sustain a desirable contract in which some individuals quit. To fix ideas while retaining simplicity, here I assume it is not. For specificity, let $z_0 = 0$ (so that limited liability prohibits firms from fining workers). Since the firm participation constraint (14) also binds with equality under any optimal contract, we can substitute it into the the limited liability constraint to get a combined constraint:

$$-v(Y^*/w) - \chi^* + \beta\tilde{u}\left(Y^* - \frac{\kappa}{F(\chi^*)}\right) \geq \beta u(0).$$

Convexity of this constraint with respect to χ^* requires that

$$1 - \beta\tilde{u}'\left(Y^* - \frac{\kappa}{F(\chi^*)}\right)\left(\frac{\kappa f(\chi^*)}{F(\chi^*)^2}\right) > 0.$$

If this inequality is violated for a given χ^*, then that χ^* cannot be sustained—intuitively, a slight reduction in χ^* raises the share who quit, reducing the utility of remaining workers, further raising the share who quit, and unraveling the contract. In the example to follow, I assume that $f(\overline{\chi})$ is sufficiently high that such unraveling occurs for all $\chi^* < \overline{\chi}$. This simplifies the contract, as the only relevant feature of the distribution $F(\chi)$ is the upper bound $\overline{\chi}$. Letting Y_0^* and Y_1^* denote the labor supply levels which satisfy self 0 and self 1's first-order conditions without any quits, the three contracting regions have a simple interpretation in this case. In the full commitment region, self 1 gets greater utility from Y_0^* than from quitting, even under the worst-case fixed cost shock $\overline{\chi}$. In the no commitment region, self 1 prefers to quit than supply even the preferred Y_1^* under the worst-cased fixed cost shock $\overline{\chi}$. In the middle limited commitment region, self 0 chooses a labor supply between Y_1^* and Y_0^* to achieve partial commitment, while partially appeasing self 1 to prevent a quit under the maximal fixed cost $\overline{\chi}$.

Homogeneous present bias. I begin with the very simplest case of homogeneous present bias, with $\beta = 0.7$. Figure 9 displays the utility from various labor supply and consumption bundles for self 0 (top panel) and self 1 (bottom panel), as a function of ability. The solid green lines plot the utility each self would derive from the labor supply which solves self 0's first-order condition. The dashed red lines show each self's utility from self 1's (lower) preferred labor supply. The horizontal dotted line in the lower panel shows the utility from quitting for an individual with maximal fixed costs $\chi = \overline{\chi}$. (I set $\overline{\chi} = 0.42$, slightly above self 1's utility without commitment at the lowest ability, in order to ensure all three contracting regions are represented.) The thin vertical lines divide the type space into the three contract regions: no commitment, limited commitment, and full commitment (left to right).

In the right-most (full commitment) region, even a self 1 with a maximal fixed cost shock $\chi = \overline{\chi}$ prefers to provide self 0's favored labor supply rather than quit—thus the first-best commitment can be sustained, individuals in this region realize utilities on the green line in equilibrium. In the center (limited commitment) region, the self 0's first-best labor supply would induce individuals with high fixed cost realizations to quit, inflating the fixed costs born by the remaining workers, and causing the contract to unravel. To prevent such an outcome, the labor supply commitment must be lowered to the point that a self 1 with $\chi = \overline{\chi}$ is just willing to remain employed, so that z satisfies

$$\beta Z^* - \kappa - \frac{(Y^*/w)^{(1+1/e)}}{1+1/e} = \overline{\chi}.$$

This modification lowers self 0's utility below the that which would be realized under full commitment, as depicted by the dot-dashed orange line in the upper figure, which plots utility as a function of ability for individuals in this region in equilibrium. Finally, in the left-most (no commitment) region, a self 1 with a sufficiently high realization of χ will prefer to quit rather than provide even his most preferred level of labor supply. That positive probability of quitting unravels any potential contract, and thus in this region individuals do not contract and instead supply their preferred labor supply in period 1 after observing their fixed cost (if they work at all).

In this setup, the labor supply elasticity is equal to e in the left and right regions and zero in the center region. The income effect is equal to zero in the left and right regions (due to quasilinear utility of consumption) and is positive in the center region. These transitions are discrete when β is assumed to be homogeneous, though a smooth U-shape and inverse U-shape emerges under two dimensional heterogeneity.

Heterogeneous present bias. I now extend the above example to a context with heterogeneous β. Suppose individuals are distributed with full support over $\left[\underline{\beta}, \overline{\beta}\right] \times [\underline{w}, \overline{w}]$. Figure 10 displays the areas of the type space corresponding to various combinations of β and w for a range of abilities (for which the scaling is immaterial) and present bias.

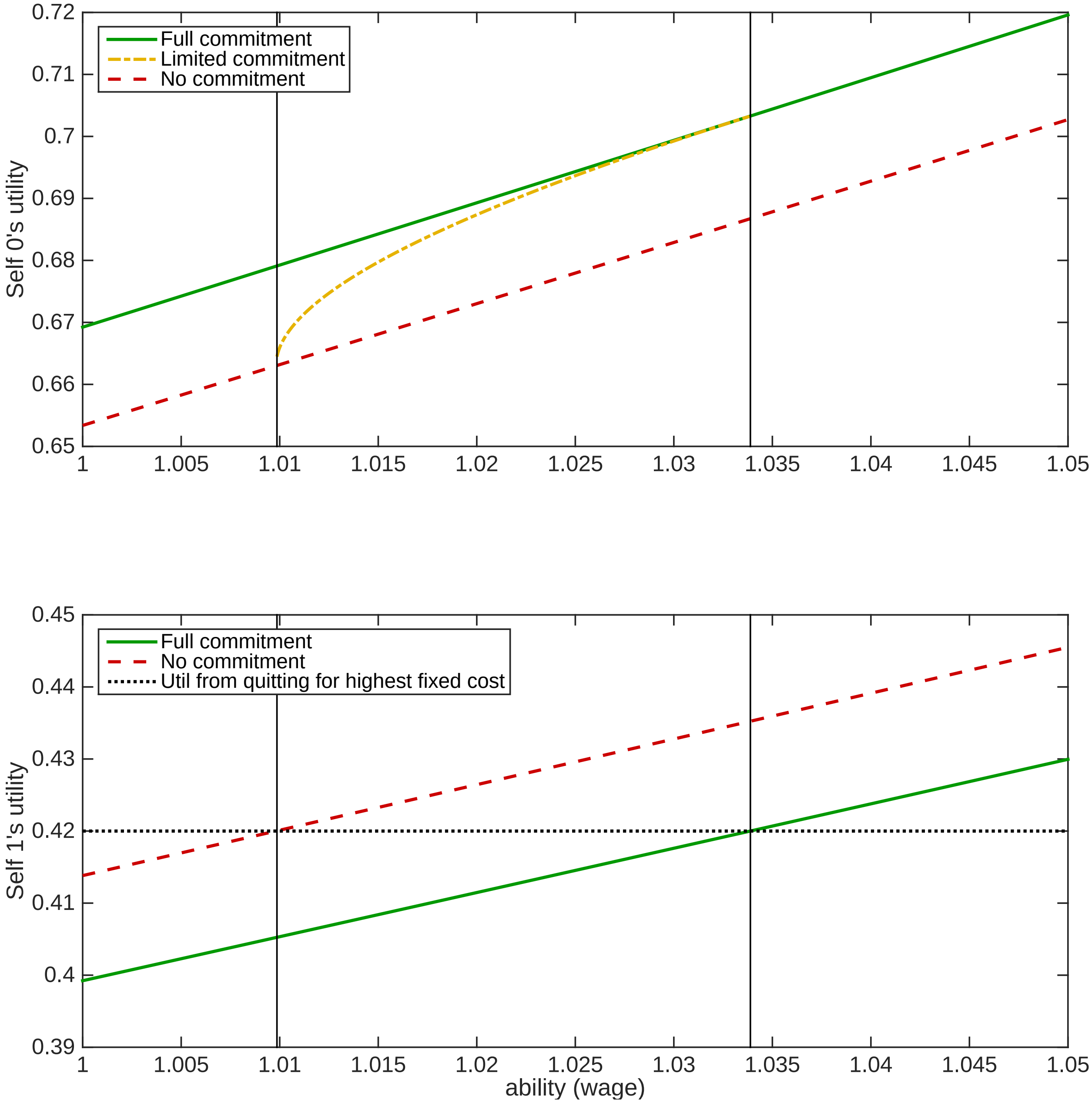

Figure 9: This figure plots the utility derived by the "long-term" self 0 (top panel) and the present biased self 1 under full commitment, no commitment, and limited commitment. Self 1's utility from the outside option of quitting, assuming the highest possible fixed cost draw $\overline{\chi}$, is shown by the dotted black line in the lower panel. Thin vertical lines divide the ability space into the three regions of employment contracts: full commitment (right), limited commitment (center), and no commitment (left).

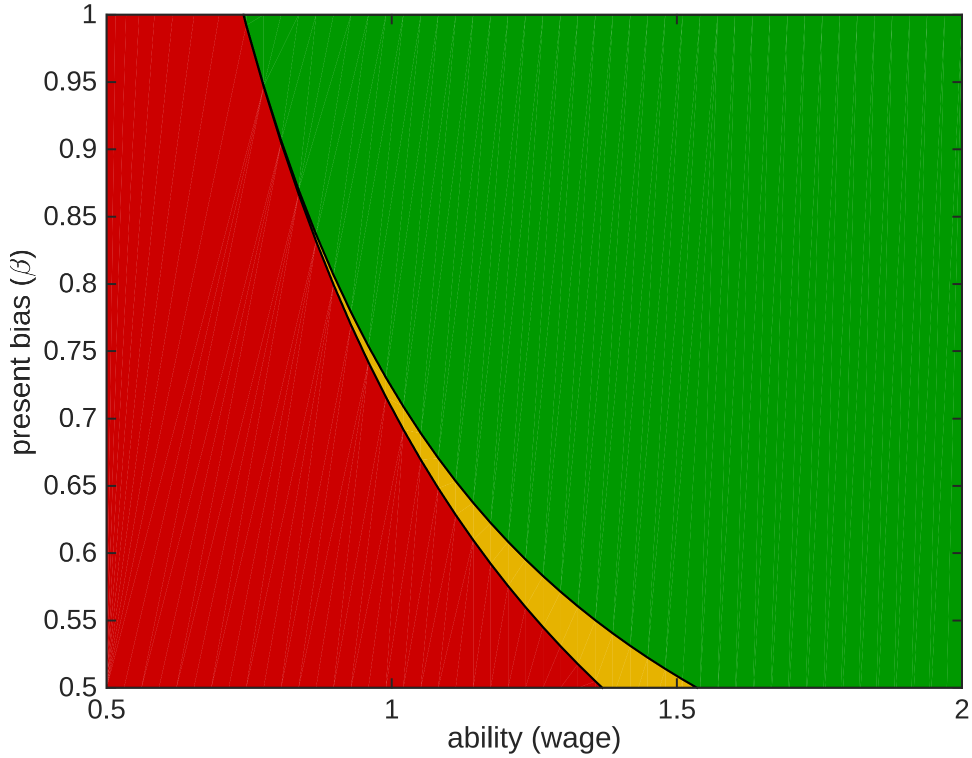

Figure 10: This figure plots the regions of the ability and present bias type space in which each contract type prevails. The upper right area represents full commitment; the middle area represents limited commitment; the lower left area represents no commitment.

The region to the lower left represents the low ability, high bias types for whom commitment is unsustainable. The middle region represents the set of types for whom limited commitment is feasible, with the participation constraint binding for $\overline{\chi}$, and the upper right region represents the set of types for whom full commitment is sustainable.

Endogenous sorting on income. The results in Figures 9–10 hold for any joint distribution over ability and present bias. However to simulate an income distribution, we must assume a distribution. For illustrative purposes, I explore the distribution of incomes and contract behavior for a uniform joint distribution over the region in Figure 10. The ability distribution can be extended upward without affecting the behavior of the income distribution above this range, for example by adding a nonparametric ability distribution calibrated to match the distribution of middle class incomes, with a Pareto tail at the top to represent the distribution of top incomes.

Figure 11 displays the resulting income densities. As noted earlier, the density is continuous. The density of incomes is fairly uniform and dispersed in the full commitment region, as the sole dimension of heterogeneity determining income variation in this region is ability. (This also explains the positive income density at the upper bound of the distribution, which consists of all individuals on the right edge of Figure 10, with $w = \overline{w}$.) In contrast, income density falls to zero at the bottom of the distribution, reflecting the variation of income with both w and β in this region—agents with earnings at the lower bound are represented by the point at the bottom left corner of the ability/bias space in Figure 10, and thus they have measure zero.

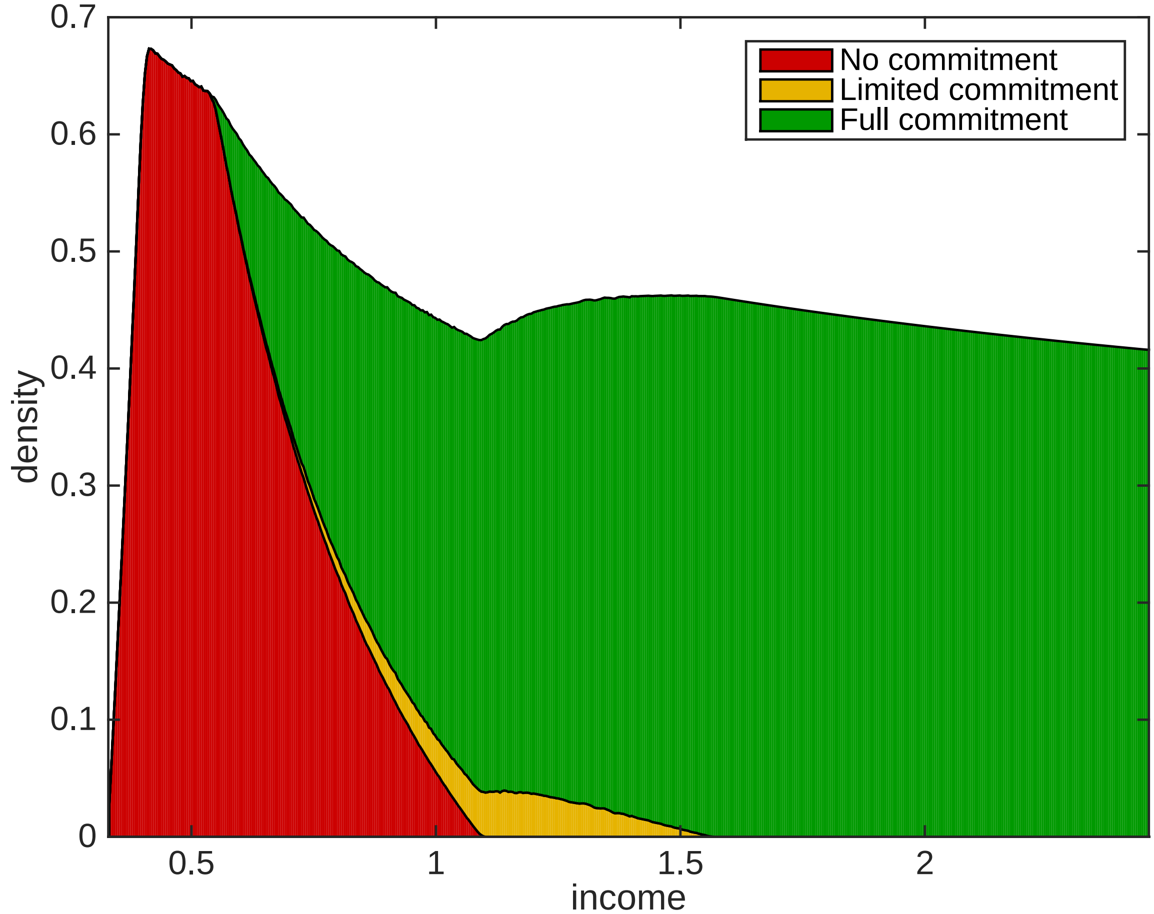

Figure 11: This figure displays the density of incomes resulting from a uniform distribution over the set of ability and bias in Figure 10. The densities of individuals falling under each contract type are shaded as in Figure 10, with full commitment dominating at high incomes, limited commitment bunched at middle incomes, and no commitment at low incomes.

Endogenous sorting on income leads to greater concentration of uncommitted workers at low incomes than was apparent over the ability distribution—note that top incomes of those without commitment is below the median, though some such individuals have above median ability (see Figure 10). This is a result of the fact that individuals with a given ability who have a sufficiently high β to sustain their first-best contract earn more than individuals who are unable to sustain a contract. Individuals with limited commitment have earnings which primarily overlap the lower tail of fully committed workers.

Appendix G Review of studies estimating structural present bias

There is a substantial experimental literature—both in the lab and in the field—testing for present bias (see DellaVigna (2009) for a review).

Choice-based estimation. As discussed in the text, choice-based approach identifies a environment in which choices are trustworthy, and assumes that similar preferences persist in alternative environments. See Bernheim and Rangel (2007) and Bernheim and Rangel (2009) for a discussion of this theoretical framework, and Chetty, Looney and Kroft (2009) for an application to nonsalient taxes.

Augenblick, Niederle and Sprenger (2013) presents a lab experiment in which student participants face a fixed amount of effort to be performed within a given period. Absent access to a commitment device, students appear wait until late in the period to perform most of the effort—but a majority of students (59%) prefer a dominated commitment contract, which leads to smoother labor effort throughout the period. Commitment demand is greater among individuals who exhibit greater present bias. Individuals without commitment exhibit an apparent discount rate of about 11% per week. Thus if individuals were time consistent (with no discounting) beyond one week, this would suggest a misoptimization wedge of 0.89—this is the estimate reported in Table 1.

Kaur, Kremer and Mullainathan (2015) measures the labor supply responses of employees in an Indian data entry center who were exposed to a number of treatments during a year-long experiment. Two findings are of particular interest. First, workers generated more output on paydays, with production rising smoothly over the weekly pay cycle as payday approached. This "payday effect" is larger than that which could be explained by a sensible time-consistent rate. Second, when given the option to choose a dominated commitment contract, in which pay is docked if an individual earns less than their pre-selected (and freely chosen) "goal amount." A substantial minority of individuals chose such contracts, and that share is larger among individuals who exhibit a more pronounced payday effect. Moreover, correlation between commitment demand and payday effect rises over time, consistent with individuals learning about their self control problems (i.e., becoming more sophisticated) over time. Kaur et al. (2015) find that the payday effect is consistent with a daily discount factor of about 5%. This suggests a discount factor of about 0.7 at a one week horizon. Their results are inconsistent with a strict β δ model of quasihyperbolic discounting, as The results of commitment demand also shed light on the degree of sophistication and time inconsistency. The authors report that the demand for dominated commitment contracts suggest that "the difference in discounting of benefits between the self that chooses the contract and the one that works [is] at least 18%." If all discounting by the contract-selecting-self is rational, and if individuals are approximately time consistent beyond a one week horizon, this suggests $\beta = 1 - 0.18 = 0.82$, which is the value reported in Table 1. Among individuals with above-median payday effects, the difference is 64%, suggesting a misoptimization wedge among these most biased individuals of 0.64.

The remaining two studies in this category do not deal with tradeoffs over labor effort, however they are of particular interest for the population they study—both measure the preferences of EITC recipients in the US. Jones and Mahajan (2015) presents a field experiment in which EITC recipients have an opportunity to sign up for short-run savings accounts yielding high interest rates, with the option to make those accounts less liquid. Meier and Sprenger (2015) presents a field experiment wherein EITC filers in Boston are given choices between intertemporal tradeoffs between monetary payments at different horizons. In both cases, subjects exhibit greater impatience at shorter horizons, suggestive of present bias. Meier and Sprenger (2015) perform a maximum likelihood estimation to calibrate the parameters of a quasi-hyperbolic discounting model. Figure 5 plots the nonparametric relationship (unreported in the paper) between the estimated β in each income quintile in their data, and average income within each quintile.[29]

The second source of evidence about present bias over labor supply comes from structural models, in which true utility is taken to be discounted exponentially. Laibson et al. (2007) use the method of simulated moments to perform a calibration using data on income, wealth, and credit use. Fang and Silverman (2009) calibrate a quasi-hyperbolic model of model of welfare takeup and labor supply using data from the NLSY.

[29] I am very grateful to Stephan Meier and Charlie Sprenger for sharing their data for this analysis.

Structural estimation. Of particular interest, DellaVigna and Paserman (2005) analyzes the relationship between measured impatience and unemployment exits, and finds evidence for present bias over labor search effort, and Paserman (2008) extends the analysis to a calibrated structural model, finding evidence consistent with quasi-hyperbolic discounting behavior. The analysis is conducted separately for the bottom quartile, the middle half, and the top quartile of wage earners. The estimated β values are $\beta = 0.40$ in the bottom quartile, $\beta = 0.48$ or $\beta = 0.81$ in the middle half (depending on assumptions about the shape of the wage distribution) and $\beta = 0.89$ in the top quartile. I use the average of the two values for the middle half, $\beta = 0.65$, and I plot the three resulting values of β against average wage earnings in the bottom quartile, middle half, and top quartile of the 2010 wage distribution in Figure 5. (Results are nearly identical if I instead multiply the previous weekly wage for each group, as reported in Paserman (2008), by 52, and convert to 2010 dollars.)