The Surprising Power of Age-Dependent Taxes

Matthew Weinzierl*

September 8, 2008

Abstract

This paper provides a new, empirically-driven application of the dynamic Mirrleesian framework by studying a feasible and potentially powerful tax reform: age-dependent labor income taxation. I show analytically how age dependence improves policy on both the intratemporal and intertemporal margins. I use detailed numerical simulations, calibrated with data from the U.S. PSID, to generate robust policy implications: age dependence (1) lowers marginal taxes on average and especially on high-income young workers, and (2) lowers average taxes on all young workers relative to older workers when private saving and borrowing are restricted. Finally, I calculate and characterize the welfare gains from age dependence. Despite its simplicity, age dependence generates a welfare gain equal to between 0.9 and 2.5 percent of aggregate annual consumption, and it captures a substantial portion of the gain from reform to the dynamic optimal policy. The gains are split approximately equally between improvements in efficiency and increases in equity. Nevertheless, when age dependence is restricted to be Pareto-improving, the welfare gain is nearly as large.

Introduction

The fundamental challenge for tax policy design, as first posed rigorously by Mirrlees (1971), is to use observable information to redistribute income to the low-skilled without discouraging work effort. The three decades following Mirrlees' initial work saw substantial progress toward meeting this challenge in his canonical model economy,1 but a recent surge of research has taken up the challenge in more complicated settings. Called the New Dynamic Public Finance, this research has focused on dynamic economies in which individuals' wages may change over their lifecycles. It has shown that dynamic optimal taxation depends, other than in special cases, on a taxpayer's detailed income history. The complexity of taxes implied by this finding has raised concerns over the practicality of the model's recommendations and has generated interest in capturing some of the gains of dynamic optimal taxation through more limited and pragmatic policy reforms. Following early, related work by Guesnerie (1977) and Feldstein (1976), these proposals are collectively called "partial reforms."2

*Harvard University: matt_weinzierl@post.harvard.edu. I am grateful to Stefania Albanesi, Robert Barro, Michael Boskin, V.V. Chari, David Cutler, Emmanuel Farhi, Martin Feldstein, Caroline Hoxby, Oleg Itskhoki, Larry Jones, Emir Kamenica, Larry Katz, Patrick Kehoe, Narayana Kocherlakota, Erzo F.P. Luttmer, Greg Mankiw, Ellen McGrattan, Thomas Mertens, Brent Neiman, Chris Phelan, Jim Poterba, Robert Shiller, Larry Summers, Aleh Tsyvinski, and Ivan Werning for their comments and suggestions.

1Saæz (2001) is the leading recent study of the standard Mirrlees model.
2The dynamic Mirrleesian optimal tax literature begins with Golosov, Kocherlakota, and Tsyvinski (2003), whose optimal policy includes history dependence. Albanesi and Sleet (2006) shows that history dependence can be replaced with dependence on current wealth if shocks to skills are i.i.d., while Golosov and Tsyvinski (2006) study optimal policy when shocks are permanent (i.e., disability shocks). For general shock processes, Kocherlakota (2005) shows that history dependence is required.
In this paper, I study a feasible but powerful partial reform that responds to changes in wages over the lifecycle: age-dependent labor income taxes. Using data from the U.S. Panel Study of Income Dynamics, I find that age dependence yields a welfare gain equivalent to between 0.9 and 2.5 percent of aggregate annual consumption, and this gain represents a substantial share of the potential gain from full reform to the dynamic optimal policy. As a new, empirically-driven application of the dynamic Mirrleesian framework that would be simple to implement in modern tax systems, age dependence is of both theoretical and practical policymaking interest.

Though I provide the first comprehensive theoretical and numerical examination of age dependence in a Mirrleesian dynamic optimal tax model, the idea of having taxes depend on age has been around for some time and was even mentioned in passing by Mirrlees (1971). The most well-known prior treatment of the subject is an elegant paper by Kremer (2002) that I discuss below. More recently, the potential for age-dependent taxation has received prominent attention from both academics and policymakers. In the 2008 Mirrlees Review, a leading summary of the current state of tax theory as it relates to policy, James Banks and Peter Diamond (2008) argue that age dependence is one of the most promising areas for the near-term reform of developed-country taxes. In existing policy, age-dependencies tend to focus on the elderly. Singapore’s public retirement savings system, the Central Provident Fund, is funded by compulsory employer and employee contributions that decline, as a share of earnings, from age 55 to retirement. Australia’s tax code includes a Senior Australian Tax Offset, which lowers taxes paid by low and middle-income Australians over the age of 65. The United States may be the site of the next age-dependent policy, as the proposal to "eliminate all income taxation of seniors making less than $50,000 per year" is one of the major policy ideas of Barack Obama, the Democratic party’s candidate for the U.S. Presidency in 2008.

In this paper, I make three specific contributions to our understanding of age-dependent taxation and dynamic optimal taxation in general.

First, I use the tools of dynamic Mirrleesian optimal tax analysis to characterize the partial reform of age-dependence analytically, deriving results along both the intratemporal and intertemporal behavioral margins that I compare to both age-independent and dynamic optimal taxation. Age dependence makes it possible to tailor the schedule of marginal labor income tax rates to the distribution of wages at each age, avoiding inefficient distortions to labor effort. Age dependence also makes it possible for tax policy to transfer resources over individuals’ lifecycles, ameliorating any limits on private saving and borrowing and shifting the intertemporal allocation of consumption closer to the optimum.

Second, I calibrate and simulate the policy models in both deterministic and stochastic settings using detailed data on heterogeneous individual wage paths. In doing so, I perform simulations that go beyond much of the dynamic optimal taxation literature, where illustrative numerical simulations are often used. This enables me to draw novel, realistic policy lessons despite the complexity of the setting, and it brings to the study of dynamic optimal taxation the empirical realism found in Saez’s (2001) study of the static Mirrlees model and Golosov and Tsyvinski’s (2006) study of optimal disability insurance. The numerical simulations of age dependence has been studied in the Ramsey tax framework. For example, Erosa and Gervais (2002) study the effect of private asset accumulation and changes in the elasticity of labor supply with age on optimal linear taxes. As a representative agent framework, the conventional Ramsey approach neglects the redistributive role of taxation. I have simulated a Ramsey version of age-dependent taxation (i.e., age-specific linear taxes and lump-sum taxes or grants) in the heterogeneous-agent model economy discussed below, and the welfare gains from age-dependence are approximately one-tenth their size in the Mirrlees approach. The Ramsey approach cannot tailor marginal taxes to variation in the distribution of wages with age.

The Mirrlees Review is the modern counterpart to the Meade Report of 1978, the influential review of taxation.

To be clear, many motivations besides the normative ones given in this paper could be driving existing proposals for age-dependent taxes.
generate specific implications for policy design. Two results largely robust across environments are that age dependence: (1) lowers marginal taxes on average and especially on high-income young workers, and (2) lowers average tax rates for young workers relative to older workers when private saving and borrowing are restricted. These results capture key features of the optimal dynamic policy.

Finally, I show that age dependence, despite being an easily-implemented partial reform, yields a large welfare gain and that this gain represents a substantial share of the potential gain from full reform to the dynamic optimal policy. This illustrates the potential for powerful policymaking lessons from dynamic optimal taxation analysis even if the full package of recommended reforms is infeasible. Importantly, I also provide a detailed decomposition of the welfare gain from age dependence. This gives us a unique look inside the machinery of Mirrleesian optimal policy, revealing exactly how age dependence generates its welfare gains. I highlight three main components. Efficiency and equity gains each account for a bit less than half of the total welfare gain, while intertemporal consumption smoothing accounts for approximately ten percent in the baseline model. Importantly, I show that the welfare gains from age dependence are nearly as large when we constrain it to be Pareto improving. As Guesnerie (1977) noted in early work on partial reforms, a particularly strong case with regard to political feasibility and normative desirability can be made for reforms that yield Pareto improvements.

As mentioned above, the most prominent prior treatment of age dependence is Kremer (2002). In a classic Mirrleesian setting, Kremer demonstrates that marginal income taxes not conditioned on age are unlikely to be optimal. He also uses evidence on some key parameters to argue that marginal tax rates should be lower for young workers, a claim supported and clarified by the theoretical analysis and detailed numerical simulations I perform below. This paper goes beyond Kremer’s analysis in three ways, each of which was mentioned in his paper as an important challenge for future research. First, Kremer’s analysis is in a static setting, while I characterize optimal age-dependent taxes in a dynamic economy. This is important because it allows me to derive the implications of factors present only in dynamic models, such as the ability of individuals to move resources intertemporally, and to show that the power of age dependence is robust to these factors. Second, while Kremer focuses on showing that first-order conditions imply age-dependence is desirable, I am able to provide a detailed description of all aspects of optimal age-dependent policy. I use panel data on individual wages to populate and simulate a model economy, yielding insights on optimal age-dependent average taxes and intertemporal distortions in addition to age-dependent schedules of marginal tax rates. Third, using my calibrated dynamic model, I calculate the large welfare gains from age dependence relative to age-independent taxes and as a share of the gain from reform to dynamic optimal taxes. While Kremer was able to argue that age dependence was likely to raise welfare, this paper’s ability to estimate that gain makes a powerful case for it as a policy reform.

As a new, empirically-driven application of the dynamic Mirrleesian framework, this paper contributes to the New Dynamic Public Finance literature. In particular, it extends the growing literature on partial reforms, such as the recent work of Farhi and Werning (2007). Farhi and Werning show how to implement, and estimate the impact of, a tax policy that satisfies the intertemporal optimality condition known as the Inverse Euler Equation in dynamic optimal tax models such as Golosov, Kocherlakota, and Tsyvinski (2003). As with this paper, they limit their policy to achieving one specific goal that might be more politically feasible than a wholesale system reform. Albanesi and Sleet (2006) take a different but related approach to capturing

6A related policy is lifetime income taxation, where an individual pays taxes or receives refunds in each year based on their (currently) expected lifetime income. The age-dependent policy studied here subsumes that policy and adds the important ability to tailor marginal taxes by age. Moreover, lifetime income taxation also uses age as an argument in the tax function, so it provides little or no advantage in simplicity relative to fully age-dependent taxation.
the value of dynamic optimal taxes through simpler policies. They focus on a dynamic economy in which
individuals’ abilities are i.i.d. across time and show that dynamic optimal taxes in that setting can be
implemented without history dependence, instead using a tax on wealth. In a similar vein, Golosov and
Tsyvinski (2006) focus on an economy with a different type of ability shock: permanent disability. They
characterize an optimal disability insurance policy that includes means-testing, an intuitively plausible result
with immediate policy relevance.

This paper also adds to a recent literature revisiting and building on the insights of Akerlof (1978),
who showed how "tagging," i.e., conditioning taxes on personal characteristics, can improve redistributive
taxation. Alesina and Ichino (2007) study the benefits of gender-dependent taxation, while Mankiw and
Weinzierl (2008) explore the case of height-dependent taxation. While age dependence may seem to be
merely another example of tagging, that is mistaken. A standard tag, such as gender, provides information
on an individual’s expected place in the distribution of lifetime income, but age reveals no information by
itself. The difference is that a standard tag divides the population into mutually exclusive groups, while the
entire population moves through all age groups. In other words, each age group has the same distribution
of lifetime incomes. To provide useful information to the tax authority, age must be combined with data
on an individual’s current income and how incomes at each age relate to lifetime income. Therefore, age
dependence cannot be fully understood with intuition from conventional tagging analysis.

More recently, two papers address components of the optimal design of age-dependent taxes. Blomquist
and Micheletto (2003) take a theoretical approach and analyze a simple dynamic economy in which all
individuals start with the same skill level when young and have one of two skill levels when older. Though
tractable, this setting prevents them from deriving results on the pattern of distortions across types at
different ages and on intertemporal distortions when there are multiple types at each age. In a substantially
more general theoretical setting that includes heterogeneity at all ages and both stochastic and deterministic
wage processes, I am able to address these two important dimensions of optimal policy. I also analytically
compare age dependence to not only age-independent policy but also the fully optimal, history-dependent
policy, a subject Blomquist and Micheletto do not address. Most important, the more complex and realistic
setting in this paper allows for a realistically-calibrated simulation of age-dependent taxes that generates
policy-relevant results and estimates the welfare gain from age dependence. Blomquist and Micheletto’s more
limited model’s results are more difficult to apply to policymaking, and they do not include any numerical
simulation of policy. Of particular value in the Blomquist and Micheletto analysis is the focus on the
potential for age dependence to be Pareto-improving. I explore that possibility in detail in Section 5 of this
paper. A second recent paper, Judd and Su (2006), takes a numerical approach. Their analysis of age
dependence is part of their broader study of multidimensional screening in optimal taxation, which they show
can substantially complicate the mathematical and computational analysis of the problem. Importantly, they
introduce to the literature numerical methods that address those complications. They apply their methods
by, among other examples, simulating age-dependent and age-independent policy for an illustrative set of
deterministic wage paths. In contrast to their simulations, I use individual-level data to calibrate a model
economy with a realistic set of deterministic and stochastic wage paths. This generates policy-relevant
results on the design of age-dependent taxes, and it allows me to estimate and decompose the potential
welfare gain from age dependence. I study the differences in age-dependent policy when individuals can and
cannot save or borrow, and I compare age-dependent and dynamic optimal taxation, topics not addressed
in their simulations.

Consistent with the small academic literature on age dependence, current age dependence in developed
countries is limited (as described above) and is often uncoordinated and unintentional. In the United States, for example, the main dependencies on age work at cross purposes. Social Security and Medicare payroll taxes are levied on all individuals but, given the pay-as-you-go structure of these programs, these taxes place a larger effective burden on the young than on the old (see Feldstein and Samwick, 1992). On the other hand, the existing disability system effectively places a larger effective burden on the old than on the young, since the latter would receive a longer string of benefits if disabled. Some transfer programs, such as the child tax credit or Earned-Income Tax Credit, are more likely to benefit the young than the old, but they also mean higher marginal rates on the young who earn enough to be in the "phase-out" region of these benefits. Finally, the deductibility of mortgage interest and charitable donations are more likely to benefit the middle-aged, for whom renting is less common and incomes are higher, than the young or old. Together, these largely unintentional age dependencies are unlikely to mimic, and may often work in the opposite direction of, the optimal age-dependent policy as studied in this paper.

The paper proceeds as follows. Section 1 describes the social planner’s problem in three policy scenarios for a baseline economy, i.e., with deterministic wage paths and no private saving or borrowing. For each policy, I analytically characterize the intratemporal and intertemporal distortions on private behavior and numerically simulate the structure of taxes. I also quantify the welfare implications of reform from the static optimal policy to age-dependent policy and the dynamic optimal policy. Section 1 serves as the reference point for the following three sections, in which I vary the assumptions about the economy to test the robustness of the baseline results. Section 2 allows individuals to save and borrow, Section 3 incorporates stochastic wage paths, and Section 4 combines these two variations. Throughout these extensions, the results from the baseline economy are largely robust. Section 5 discusses a range of specific topics that fall outside the main analysis of the paper, and Section 6 concludes. A Technical Appendix contains the proofs of the propositions and supplementary material mentioned in the text.

1 Baseline economy

In this section I analyze age-dependent labor income taxes for a baseline economy characterized by two simplifying assumptions. First, individuals cannot transfer resources across periods: that is, they can neither save nor borrow. Second, each individual’s lifetime wage path is deterministic, so that each individual knows in advance the exact path of wages it will have over its lifetime (i.e., there are no stochastic shocks to wages). These simplifying assumptions allow for cleaner analytical results, but I generalize the model to relax them in later sections of the paper. I start by setting up the economy and then specifying the social planner’s problem in three policy scenarios.

1.1 Setup

All individuals live and work for \(T\) periods, indexed by \(t = \{1, 2, ..., T\}\), and are members of the same generation.\(^8\) Individuals are heterogeneous in their ability to earn income over their lifetimes. This ability

\(^7\)I assume an exogenous date of entry into the labor market. This matches the quantitative analyses below, which consider age-dependent taxes designed to have minimal effects on incentives to obtain higher education or training. Age dependence could generate even larger welfare gains if it were designed to operate effectively on these incentives.

\(^8\)In the Technical Appendix, I discuss how the analysis generalizes largely unchanged to a setting with overlapping generations so long as no intergenerational transfers are allowed. In the case of this plausible restriction, each generation is effectively isolated. If intergenerational transfers were allowed, tax policy would have a new purpose that may be achievable without age dependence, so the calculation of the gains from age dependence would be less clear. In the Technical Appendix, I also discuss the relationship of calendar-year dependence to age dependence.
comes in $I$ types, indexed by $i = \{1, 2, ..., I\}$, with probabilities $\pi^i$ so that $\sum_{i=1}^{I} \pi^i = 1$. At each age $t$, an individual of type $i$ can earn a wage $w^i_t$ for each unit of its labor effort, and each individual knows its full lifetime path of wages $\{w^i_t\}_{t=1}^{T}$ at time $t = 1$. Wages are not publicly observable. I refer to the present value of wages for type $i$ as type $i$’s lifetime income-earning ability, defined as $\sum_{t=1}^{T} \frac{w^i_t}{R^t}$, where $R$ is the exogenous gross rate of return. Types are sorted so that $i = I$ refers to the type with the highest lifetime income-earning ability.

Production is structured as follows. Labor income $y$ is the product of the wage and labor effort $l$, so $y = wl$. There is no physical capital in the economy.

All individuals have the same separable preferences over consumption $c$ and labor effort $l$, where $l = \frac{y}{w}$. The utility $U^i_t$ for individual $i$ of age $t$ is

$$U^i_t(c, y) = u(c) - v\left(\frac{y}{w^i_t}\right),$$

where I assume $u'(\cdot) > 0$, $u''(\cdot) < 0$, $v'(\cdot) > 0$, $v''(\cdot) > 0$. For an individual $i$, lifetime utility $V^i$ is the discounted sum of its utility at each age:

$$V^i = \sum_{t=1}^{T} \beta^{t-1} U^i_t(c, y),$$

where individuals discount utility flows with the factor $\beta$. Social welfare $W$ is a weighted sum of individual lifetime utilities:

$$W = \sum_{i=1}^{I} \pi^i \alpha^i V^i,$$

where $\alpha^i$ indicates a scalar Pareto weight on individual $i$. In general, the form of (3) allows us to consider any point along the Pareto frontier.

### 1.2 Social planner’s problem in three policy scenarios

Now I derive optimal taxes in three policy scenarios using the techniques of modern dynamic optimal tax analysis. In this approach, the tax problem is recast as a problem for a fictitious social planner that uses a direct mechanism to allocate resources (see Golosov, Tsyvinski, and Werning, 2006 for a review).\(^9\)

The social planner maximizes social welfare (3) by offering a menu of income and consumption pairs to individuals. Individuals choose optimally from the menu, earn the assigned income, and receive the assigned consumption. Knowing this, the planner designs its menu of $\{c, y\}$ pairs intending each pair to be chosen by a specific individual. Because individuals differ in their lifetime income-earning ability and age, I write $c^i_t$ and $y^i_t$ for the pair intended for the individual of type $i$ and age $t$.

The planner maximizes social welfare subject to two types of constraints: a feasibility constraint and incentive constraints. The feasibility constraint is:

$$\sum_{i=1}^{I} \pi^i \sum_{t=1}^{T} R^{T-t} (y^i_t - c^i_t) = 0.$$  \(^{(4)}\)

\(^9\)In the Technical Appendix, I show how the optimal allocations for each policy scenario derived below using a direct mechanism can be implemented using nonlinear tax functions that depend on income, income and age, or income history and age, respectively. The formal approach follows Kocherlakota (2005).
which says that the lifetime paths of income must fund the lifetime paths of consumption across all types.\footnote{Note that taxation is purely redistributive. A positive net revenue requirement would imply larger tax distortions on average, likely increasing the welfare gain from age dependence calculated below.} The planner can transfer resources across time and earn or pay the gross rate $R$. I assume $\beta R = 1$ for simplicity.

Incentive constraints reflect that individuals choose from the planner’s menu of $\{c, y\}$ pairs to maximize their utility. In this approach, these constraints take the form of inequalities ensuring that each individual chooses the allocation of $c_i^t$ and $y_i^t$ that the planner intended.\footnote{These incentive constraints reflect this approach’s application of the Revelation Principle, by which we can restrict attention to incentive-compatible direct mechanisms, i.e., where individuals reveal their true types to the planner.}

Importantly, variations in the set of incentive constraints allow us to succinctly distinguish the planner’s problem in three policy scenarios: Static Mirrlees, Partial Reform, and Full Optimum. I now state these planner’s problems formally and discuss the differences between them.

The Static Mirrlees planner in the baseline model solves the following problem:

**Problem 1** (Static Mirrlees: Age-Independent)

Choose $\{c_i^t, y_i^t\}_{i=1,t=1}^{I,T}$ to maximize (3) subject to the feasibility constraint (4) and the incentive constraints

\[
\beta^{t-1} \left( u(c_i^t) - v(\frac{y_i^t}{w_i^t}) \right) \geq \beta^{t-1} \left( u(c_j^t) - v(\frac{y_j^t}{w_i^t}) \right)
\]

for all $i, j \in \{1, 2, ..., I\}$ and $t, s \in \{1, 2, ..., T\}$.

These incentive constraints mean that the Static Mirrlees planner must guarantee that each individual of type $i$ and age $t$ chooses the allocation intended for it over that intended for any other individual of type $j$ and any age $s$. To see this, note that each side of the inequality (5) equals period utility for an individual of type $i$ and age $t$. The left-hand side is the utility this individual obtains by earning $y_i^t$ and consuming $c_i^t$, while the right-hand side is the utility it obtains by earning $y_j^t$ and consuming $c_j^t$. The inequality guarantees that this individual weakly prefers its $\{c, y\}$ allocation.\footnote{This is different from the requirement that two individuals with the same wage but different ages receive the same allocations of $c$ and $y$. The latter is a stronger condition and restricts the Static Mirrlees planner more than is justified. Age may affect an individual’s optimal choice of income and consumption, even if its wage is unchanged, and the Static Mirrlees planner can take advantage of this without making taxes age-dependent.} I denote the multipliers on these constraints with $\{\mu_{ij}^t\}_{s,t}^{i,j}$, where $\mu_{ij}^t$ corresponds to the constraint preventing individual $i$ of age $t$ from preferring the allocation intended for individual $j$ of age $s$.

Note that the Static Mirrlees policy is not designed to match the detailed structure of existing tax policy on labor income. Rather, it is the optimal policy constrained by two characteristics of existing tax policy: age independence and history independence. Our comparison of the Partial Reform to this Static Mirrlees policy rather than to existing tax policies allows us to isolate the potential for age dependence to generate welfare gains by itself, relative to an age-neutral benchmark.\footnote{In addition to the simulations of optimal policy presented below, I have simulated a policy approximating the current U.S. tax system. Using the data and parameterization described in the paper, I apply the 1999 U.S. income tax schedule to labor earnings and simulate individuals’ behavior. The welfare gain from that policy to the Static Mirrlees policy is equivalent to a 6.9 percent increase in aggregate annual consumption.}

The Partial Reform planner in the baseline model solves the following problem:

**Problem 2** (Partial Reform: Age-Dependent)
Choose \( \{c^i_t, y^i_t\}_{i=1, t=1}^{I, T} \) to maximize (3) subject to the feasibility constraint (4) and the incentive constraints

\[
\beta^{t-1} \left( u \left( c^i_t \right) - v \left( \frac{y^i_t}{w^i_t} \right) \right) \geq \beta^{t-1} \left( u \left( c^j_t \right) - v \left( \frac{y^j_t}{w^j_t} \right) \right) \tag{6}
\]

for all \( i, j \in \{1, 2, ..., I\} \) and \( t \in \{1, 2, ..., T\} \)

These incentive constraints reflect the Partial Reform planner’s ability to restrict individuals to age-specific allocations. Now, the planner must guarantee that each individual \( i \) of age \( t \) chooses the allocation intended for it over that intended for any other individual \( j \) of the same age \( t \). To see this, notice that the right-hand side of (6) depends on \( c^i_t \) and \( y^i_t \), so that both sides of the inequality are specific to age \( t \) (compare this to the Static Mirrlees planner, where the right-hand side depended on \( c^i_s \) and \( y^i_s \)). Formally, the set of constraints (6) is a subset of (5). This makes the set of incentive constraints in the Partial Reform planner’s problem weakly easier to satisfy than the set in the Static Mirrlees planner’s problem. I denote the multipliers on these constraints with \( \{\mu^i_{j} \}_{i}^{j} \), where \( \mu^i_{j} \) corresponds to the constraint preventing individual \( i \) of age \( t \) from preferring the allocation intended for individual \( j \) of age \( t \).

Finally, the Full Optimum planner in the baseline model solves the following problem:

**Problem 3 (Full Optimum: Age-Dependent and History-Dependent)**

Choose \( \{c^i_t, y^i_t\}_{i=1, t=1}^{I, T} \) to maximize (3) subject to the feasibility constraint (4) and the incentive constraints

\[
\sum_{t=1}^{T} \beta^{t-1} \left( u \left( c^i_t \right) - v \left( \frac{y^i_t}{w^i_t} \right) \right) \geq \sum_{t=1}^{T} \beta^{t-1} \left( u \left( c^j_t \right) - v \left( \frac{y^j_t}{w^j_t} \right) \right) \tag{7}
\]

for all \( i, j \in \{1, 2, ..., I\} \).

These incentive constraints reflect the Full Optimum planner’s ability to make, and commit to, history-dependent allocations.\(^{14} \) History dependence allows the planner to hold an individual to the lifetime path of allocations intended for a single type at all ages. Thus, the Full Optimum planner must guarantee only that each individual \( i \) chooses the lifetime path of allocations intended for it over that intended for any other type \( j \). This is apparent from (7) in that each side of the inequality is a discounted sum of period utilities over individual \( i \)’s lifetime. The left-hand side is \( i \)’s lifetime utility if it chooses its intended allocations \( (c^i_t, y^i_t) \) at each age \( t \), while the right-hand side is \( i \)’s lifetime utility from claiming the allocations intended for type \( j \) at each age. I denote the multipliers on these constraints with \( \{\mu^i_{j} \}_{i}^{j} \), where \( \mu^i_{j} \) corresponds to the constraint preventing individual \( i \) from preferring the lifetime allocation intended for individual \( j \).

Using history dependence to satisfy incentives on a lifetime basis can be a powerful tool for the planner. For example, suppose the planner wants to give individual \( i \) a generous allocation later in life in exchange for a "bad" allocation early in life. In the Full Optimum, the planner can offer that path of allocations to the individual because it can make later allocations dependent on earlier ones. In the Static Mirrlees or Partial Reform scenarios, such a path is not sustainable. In those scenarios, individuals know that the planner cannot reward early sacrifice because it cannot use history dependence, so they will not accept the bad allocation early in life.

\(^{14}\)The assumption that the planner can commit to a path of allocations is standard in the dynamic optimal tax literature. Bisin and Rampini (2005) study the impacts of relaxing that assumption.
1.3 Analytical results

Now, I compare the allocations chosen by the social planner in the Static Mirrlees, Partial Reform, and Full Optimum policies along two margins: the intratemporal margin between consumption and leisure and the intertemporal margin between consumption in one period and the next. I evaluate average taxes, a key dimension along which the three policy scenarios differ, in the section using numerical simulations.

1.3.1 Intratemporal distortions

First, I compare the distortions to individuals’ choices of how much income to earn: i.e., distortions on their marginal choices between consumption and leisure.

Definition 1 (Intratemporal Distortion) The intratemporal distortion for an individual of type $i$ and age $t$ is denoted $\tau(i, t)$ and equals

$$\tau(i, t) = 1 - \frac{v'(\frac{w_i^c}{w_i^T})}{w_i^c w_i^T v'(c_i^t)}.$$  \hfill (8)

Denote $\tau^{SM}(i, t)$, $\tau^{PR}(i, t)$, and $\tau^{FO}(i, t)$ as the intratemporal distortions for an individual of type $i$ and age $t$ in the solutions to the Static Mirrlees (SM), Partial Reform (PR), and Full Optimum (FO) planner’s problems.

In expression (8), positive $\tau(i, t)$ distorts the individual’s choice away from work (and consumption) and toward leisure. If $\tau(i, t) = 0$, the individual sets the marginal utility from an extra unit of consumption equal to the marginal disutility of earning it, so there is no distortion on this margin.

If wage paths were constant for each individual over the lifecycle, $\tau(i, t)$ would be the same for individual $i$ at all ages. Moreover, each policy (Static Mirrlees, Partial Reform, and Full Optimum) would use the same set of $\tau(i, t)$ in this economy, as all ages would be identical.\(^{15}\)

In reality, wage paths are far from constant, and optimal policy will vary intratemporal distortions in response if it is able to do so. As in most Mirrleesian analyses, characterizing intertemporal distortions analytically is difficult, and the general expressions for $\tau(i, t)$ are not available in closed form. Nevertheless, we can use expressions for them to illustrate the key lessons about age-dependent marginal taxes. To simplify the results, I assume disutility takes the isoelastic form

$$v\left(\frac{y_i^c}{w_i^T}\right) = \frac{1}{\sigma} \left(\frac{y_i^c}{w_i^T}\right)^\sigma.$$  \hfill (9)

where the parameter $\frac{1}{\sigma - 1}$ gives the constant-consumption elasticity of labor supply.

In the Static Mirrlees scenario, the intratemporal distortion on worker of type $i$ and age $t$ is:

$$\tau^{SM}(i, t) = \frac{\alpha^i \pi^i}{\alpha^i \pi^i + \sum_{s=1}^T \sum_{j=1}^J \left(1 - \left(\frac{w_i^c}{w_j^s}\right)^\sigma\right) \beta^{s-t} \mu^{ij}_{s|t}}.$$  \hfill (10)

where, as stated after (5), $\mu^{ij}_{s|t}$ is the multiplier on the incentive constraint preventing individual $i$ of age $t$ from claiming the allocation of any other individual $i$ of age $s$.\(^{16}\)

\(^{15}\)I prove these results in an earlier version of this paper, Weinzierl (2008), Proposition 1.

\(^{16}\)As Judd and Su (2006) point out, it is possible that optimal policies in these settings include pooling multiple types of individuals at the same allocation. In that case, the values of the incentive constraints’ multipliers may not be uniquely defined.
In the Partial Reform scenario, it is:

$$\tau^{PR} (i, t) = \frac{\sum_{j=1}^{I} \left(1 - \left(\frac{w^i_j}{w^j_i}\right)^\sigma \right)}{\alpha^i \pi^i + \sum_{j=1}^{I} \mu^j_i \mu^i_j} \mu^i_j$$

(11)

where $\mu^i_j$ is the multiplier on the incentive constraint preventing individual $i$ of age $t$ from claiming the allocation of any other individual $j$ of the same age $t$.

Finally, in the Full Optimum scenario, it is:

$$\tau^{FO} (i, t) = \frac{\sum_{j=1}^{I} \left(1 - \left(\frac{w^i_j}{w^j_i}\right)^\sigma \right)}{\alpha^i \pi^i + \sum_{j=1}^{I} \mu^j_i \mu^i_j} \mu^i_j$$

(12)

where $\mu^i_j$ is the multiplier on the incentive constraint preventing individual $i$ from claiming the lifetime allocation of any other individual $j$.

By comparing these expressions we can learn a key lesson: age-dependent taxes allow policymakers to tailor intratemporal distortions to the wage distribution at each age. In contrast, optimal age-independent taxes distortions must be based on the distribution of wages across all ages. To see this contrast formally, note that expression (11) depends on wage ratios within age $t$, while expression (10) depends on wage ratios across ages $t$ and $s$. The multipliers $\mu^i_{j|a}$, $\mu^j_i$, and $\mu^i_{ij}$ introduce some complexity into these expressions, but the distinction remains.

The clearest example of age-dependent taxes being tailored to age-specific wage distributions builds on the most well-known result from static Mirrleesian tax analysis: the top earner in the economy should face no intratemporal distortion.\(^\text{17}\) The intuition for the classic result is that an intratemporal distortion has both a cost and a benefit. The cost is that it causes individuals to work and consume differently than they would without taxes, leading to either lower utility or less efficiently-provided utility for these individuals. The benefit is that it enables the planner to collect more tax revenue from higher earners, increasing the extent of redistribution. At the top of income distribution, this benefit is zero (there are no higher earners from whom to collect more tax revenue). Thus, a distortion on the top earner solely discourages effort, and it is avoided in the optimal policy.

The following proposition describes how this classic result applies to a dynamic economy under this paper’s three policy scenarios:

**Proposition 1 (Top Marginal Distortion)** In the baseline economy, for

1. if $w^i_j \geq w^i_j$ for all $j \in \{1, 2, ..., I\}$, then $\tau^{PR} (i, t) \leq 0$ and $\tau^{FO} (i, t) \leq 0$,

2. if $w^i_j \geq w^i_j$ for all $j \in \{1, 2, ..., I\}$ and for all $t \in \{1, 2, ..., T\}$, and if $\alpha^i \leq \alpha^j$ for all $j \in \{1, 2, ..., I\}$, then $\tau^{PR} (i, t) = 0$ and $\tau^{FO} (i, t) = 0$ for all $t \in \{1, 2, ..., T\}$, and $\tau^{SM} (i, t) = 0$ for $t$ such that $w^i_s \geq w^i_s$ for all $s \in \{1, 2, ..., T\}$.

These expressions implicitly assume that some values for these has been chosen, but the discussion of the expressions does not presume anything about those values.

\(^{17}\)Diamond (1998) and Saez (2002) show that this result depends on the shape of the wage distribution. With a bounded wage distribution such as that used throughout the paper, the zero top rate result always holds. In the Technical Appendix, I show that the results presented here have analogues in a model with a wage distribution more similar to that which they use. While the top rates are not zero at any age for any policy, rates for the high-skilled vary by age, reflecting the different shapes of the skill distributions at each age.
Proof. In Technical Appendix.

The first part of this proposition states that the highest wage earner at each age faces a nonpositive, even negative, intratemporal distortion whenever the social planner can condition taxes on age.\(^\text{18}\) Intuitively, suppose an individual has the highest wage within its current age but also has a low lifetime income-earning potential. An example might be an entertainer or athlete whose earnings power is temporarily high at a young age. The planner wants to assign this individual both high income and high consumption: the former because it is a productive worker, and the latter because its welfare weight is relatively large. A negative marginal distortion makes this possible. With it, the planner can levy high average taxes on individuals who earn less at the current age but more over their lifetimes, for example consultants or lawyers, while using the negative distortion to reduce the tax burden on the top current earner. The lower earners will not be willing to earn enough income to qualify for the negative distortion, so the planner is able to target its resources.

The second part of this proposition states that an individual who is the highest wage earner at all ages faces no intratemporal distortion at any age in the Partial Reform and Full Optimum scenarios, but only at its peak-earnings age in the Static Mirrlees scenario. To see why it holds, consider an individual \(i^*\) with the highest wage at all ages. As mentioned above, the potential benefit of a positive marginal distortion on any individual is that it makes possible the collection of more tax revenue from higher earners. If there are no higher earners than \(i^*\) at each age, positive distortions on \(i^*\) have costs but no benefits, so they are avoided by the planners with access to age-dependent taxes. The Static Mirrlees planner, on the other hand, faces a more difficult problem. While no other individuals of the same age earn more than \(i^*\), some individual of a different age (perhaps \(i^*\) itself) earns more than \(i^*\) for each age except the age at which \(i^*\)’s earnings peak. Thus, in order to collect revenue from the highest earners across all age groups, the planner will use distortions on \(i^*\) at all but its peak-earnings age. A simple numerical illustration of this is provided in the Technical Appendix for the interested reader.

Though the second part of the proposition is about a special topic, the top marginal distortion, it highlights the fundamental limitation of the Static Mirrlees policy: the inability to hold individuals to age-specific tax schedules. The effects of this limitation ripple throughout the wage distribution, and we will quantify these effects and their impact on social welfare in the numerical simulations below.

In those simulations, we will also be able to address some natural questions about the structure of intratemporal distortions in an age-dependent tax policy that are difficult to answer analytically. In particular, does an individual face rising or falling intratemporal distortions over its lifecycle? If two individuals earn the same current income at different ages, who should face the higher distortion on that income? As we will see, the answers to these important questions depend critically on the details of the economy and the structure of wages. Before turning to the simulations, however, I discuss a second set of analytical results, these on the intertemporal consumption margin.

### 1.3.2 Intertemporal Distortions

In this section, I analytically characterize the intertemporal consumption margin in an age-dependent policy and compare it to well-known results from the dynamic optimal tax literature. I show that age-dependent policy satisfies a condition that improves on age-independent policy but falls short of the full optimum. The optimality of a negative top distortion has also been suggested by Judd and Su (2006), but for a different reason. In their model with multiple dimensions of heterogeneity, the interaction of wage and labor supply elasticity differences can justify negative top distortions. As noted in the Introduction, Judd and Su also perform an illustrative simulation of age-dependent taxation.
recent development of a dynamic Mirrlees literature has highlighted this margin because optimal distortions to it, most prominently characterized by Golosov, Kocherlakota, and Tsyvinski (2003), have renewed interest in the taxation of capital after a long period during which the Chamley-Judd result of zero optimal capital taxation held sway.

Consider an individual’s problem of maximizing lifetime utility (2) given a wage path \( \{w^t_i\}_{t=1}^T \) and the lifetime budget constraint \( \sum_{t=1}^T R^{T-t} (y^t_i - c^t_i) = 0 \). Continue to assume \( \beta R = 1 \). This individual’s optimal choice of consumption satisfies, for each \((t, t+1)\) pair:

\[
u' (c^t_i) = u' (c^{t+1}_i). \tag{13}\]

This is the familiar intertemporal Euler condition that sets the marginal utility from consumption equal across periods. Expression (13) represents an undistorted intertemporal margin. Recall that individuals cannot save or borrow in the baseline economy. Thus, when studying this margin, we are interested in the extent to which the planners’ chosen allocations distort allocations away from (13).

If wage paths were constant, then each policy (Static Mirrlees, Partial Reform, and Full Optimum) would satisfy (13).\(^{19}\) Intuitively, if each age is a replica of the next, the allocations to each individual will be the same at each age.

In more realistic settings, with wage paths that are not constant over the lifecycle, we can show how each policy responds along the intertemporal margin.

The Full Optimum planner’s allocations satisfy, for all \( i \in \{1, 2, \ldots, I\} \) and \( t, t+1 \in \{1, 2, \ldots, T\} \):

\[
u' (c^t_i) = u' (c^{t+1}_i), \tag{14}\]

a classic Atkinson and Stiglitz (1976) result: i.e., the Full Optimum policy does not distort intertemporal allocations.\(^{20}\) This result depends on the Full Optimum planner’s ability to use history-dependent allocations, as is made clear by the contrast between expression (14) and the result for the Partial Reform planner to which I now turn.

The Partial Reform planner’s allocations satisfy, for individual \( i \) and ages \( t, t+1 \):

\[
u' (c^t_i) = u' (c^{t+1}_i). \tag{15}\]

The ratio in parentheses in (15) is generally different from one, implying that the Partial Reform planner imposes a distortion on the intertemporal margin. To see this, recall that \( \mu^{ji}_t \) and \( \mu^{ji}_{t+1} \) are the multipliers on the incentive constraints preventing type \( i \) from claiming type \( j \)’s allocation at ages \( t \) and \( t + 1 \), respectively. Unless \( \mu^{ji}_t = \mu^{ji}_{t+1} \) for all \( i, j, \) and \( t \), the ratio in parentheses in (15) is not equal to one. Intuitively, whenever the incentive problems facing the planner differ across ages, the Partial Reform planner generally fails to satisfy the intertemporal Euler equation. The Partial Reform planner must satisfy incentives at each age, so changing wage distributions make the allocation for an individual differ across ages, violating the intertemporal Euler equation.

The ability to satisfy the standard Euler equation and provide smooth consumption is a substantial

\(^{19}\) I show this in an earlier version of this paper, Weinzierl (2008), as Proposition 3. Werning (2007a) proves a similar result for the optimal dynamic policy.

\(^{20}\) Note that wages are deterministic, so there is no reason for an intertemporal distortion along the lines of Rogerson (1985) or Golosov, Kocherlakota, and Tsyvinski (2003). In Sections 3 and 4, I analyze generalizations of the baseline model where wages are stochastic.
advantage for the Full Optimum planner relative to the Partial Reform (or Static Mirrlees\textsuperscript{21}) planner. Nevertheless, the Partial Reform allocations do improve on the Static Mirrlees allocations, as they satisfy what I will call the "Symmetric Inverse Euler Equation:"

**Proposition 2 (Symmetric Inverse Euler)** Let the lifetime utility function for all $i \in \{1, 2, ..., I\}$ be defined by (2). Then, the solution to the Partial Reform planner’s problem satisfies:

$$
\sum_{i=1}^{I} \pi_i \frac{u'(c(i, t))}{w'(c(i, t))} = \sum_{i=1}^{I} \pi_i \frac{u'(c(i, t + 1))}{w'(c(i, t + 1))},
$$

for any $t, t + 1 \in \{1, 2, ..., T\}$.

**Proof.** In Technical Appendix. ■

This "Symmetric Inverse Euler Equation" guarantees that resources are being allocated efficiently between age groups, as it equalizes across ages the cost (in consumption) of increasing welfare. Though not as powerful a restriction as the intertemporal Euler equation, the Symmetric Inverse Euler Equation is nevertheless an achievement of Partial Reform that the Static Mirrlees planner cannot replicate. Because it cannot restrict individuals to age-specific tax schedules, the Static Mirrlees planner cannot make these efficient transfers across ages.

### 1.3.3 Summary

The analytical results above make clear that an age-dependent policy differs substantially from an age-independent one and resembles in many important ways a fully optimal policy. They show how theoretical results for age-dependent policy in a dynamic economy with lifecycle wage paths connect to and extend prominent results from the static and dynamic optimal tax literature. As with most Mirrleesian analyses, however, these results can pin down only select characteristics of tax policy. To provide a more general characterization of policy, I turn to numerical simulations.

### 1.4 Numerical results

In this section, I calibrate the dynamic optimal tax problems specified above to detailed individual wage data from the U.S. PSID, simulate and characterize policy, and quantify the welfare impacts of reform.

#### 1.4.1 Data

Simulating the model requires representative lifetime wage paths $i = \{1, 2, ..., I\}$. The construction of the dataset required to calculate these paths can be divided into four steps.\textsuperscript{22} First, I focus on household heads from the U.S. PSID core sample for the years 1968-2001 and collect data on their income, hours worked, age, race, gender, and education for each year they are the heads of households. Second, I calculate reported real wages for each observation by dividing reported labor income by reported hours (a potentially noisy measure of the wage but the best one available) and inflating or deflating the data with the CPI to put all

\textsuperscript{21}I omit the Static Mirrlees planner’s intertemporal result for brevity, but it is an intuitive modification of (15).

\textsuperscript{22}The process detailed in this paragraph follows Fullerton and Rogers (1993) and recent updates of their work such as Altig, Auerbach, Kotlikoff, Smetters, and Walliser (2001) and Diamond and Tung (2006). I thank John Diamond and Joyce Tung for providing helpful advice on the construction of the dataset.
wages in 1999 dollars. Third, I remove potentially problematic observations by eliminating all those for which reported annual hours were less than 500 or greater than 5,824, for which reported labor income was zero but hours were positive, or for which the nominal wage implied by earnings and hours was less than half the applicable minimum wage in that calendar year. Fourth, I limit the sample to the ages 30 through 59, as that is the range over which the numerical simulations below will be run. After these adjustments, the dataset contains approximately 95,000 observations on just over 10,000 individuals with an average of 9.5 years observed per person.

Once the dataset is assembled, the next task is to identify $i = 1, 2, ..., I$ wage paths that best represent the sampled individuals’ paths. I do so using latent class analysis to estimate $I$ sets of parameters, one for each latent type, for the regression:

$$\ln w_t = \alpha + \beta_1 (age_t) + \beta_2 (age_t)^2 + \beta_3 (age_t)^3 + \beta_4 (age_t)^4$$

where $w_t$ and $age_t$ are the vectors of wages and ages for all individuals at each age $t$. The index $i = 1, 2, ..., I$ denotes the latent types to which each individual is assigned. The parameter $\alpha$ is a vector of individual fixed effects, while the parameter vector $\beta = \{\beta_1, \beta_2, \beta_3, \beta_4\}^I_{i=1}$ takes on $I$ sets of values, one set for each latent type. Latent class analysis uses Maximum Likelihood methods to find the optimal $\beta$, and I use the program LatentGold to identify these paths. I then use the regression results to predict $I$ wage paths for use in the simulations.

Due to computational speed considerations, I limit the size of the simulation by setting $I = 10$ and by grouping wages into three main decades of work life. Thus, $i = 1, 2, ..., I$ and $t = 1, 2, 3$. The results are shown in Table 1 along with demographic data describing the income groups.

**Table 1: Data for simulation of baseline model**

<table>
<thead>
<tr>
<th>Wage paths</th>
<th>Age range</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average wage by age range ($1999)</td>
<td>30-39</td>
<td>6.50</td>
<td>8.55</td>
<td>11.53</td>
<td>12.42</td>
<td>15.06</td>
<td>19.23</td>
<td>20.97</td>
<td>23.89</td>
<td>31.52</td>
<td>42.95</td>
</tr>
<tr>
<td>40-49</td>
<td>6.38</td>
<td>9.06</td>
<td>12.52</td>
<td>12.53</td>
<td>17.05</td>
<td>21.74</td>
<td>23.97</td>
<td>27.97</td>
<td>39.08</td>
<td>63.65</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>6.07</td>
<td>9.26</td>
<td>13.24</td>
<td>11.77</td>
<td>17.46</td>
<td>23.07</td>
<td>21.44</td>
<td>29.71</td>
<td>40.83</td>
<td>71.44</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Descriptive data</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population proportion</td>
<td>0.14</td>
<td>0.11</td>
<td>0.11</td>
<td>0.16</td>
<td>0.12</td>
<td>0.11</td>
<td>0.09</td>
<td>0.08</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>Proportion race=white</td>
<td>0.37</td>
<td>0.42</td>
<td>0.55</td>
<td>0.65</td>
<td>0.64</td>
<td>0.75</td>
<td>0.79</td>
<td>0.81</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>Proportion with college degree</td>
<td>0.04</td>
<td>0.03</td>
<td>0.06</td>
<td>0.14</td>
<td>0.12</td>
<td>0.21</td>
<td>0.38</td>
<td>0.36</td>
<td>0.63</td>
<td>0.76</td>
</tr>
<tr>
<td>Proportion gender=male</td>
<td>0.55</td>
<td>0.60</td>
<td>0.72</td>
<td>0.85</td>
<td>0.83</td>
<td>0.89</td>
<td>0.93</td>
<td>0.91</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>Pareto weight (calculated)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Source: PSID core sample household heads 1968-2001; author’s calculations as described in paper.

Figure 1 plots these wage paths.

---

23 When using an empirical income distribution to infer the distribution of skills and simulate optimal taxes, it is important to back out the effects of the current tax system on income as in Saez (2001). With data on income and hours, it is possible to calculate wages directly, instead.  
24 Detail on latent class methods and Latent Gold in particular can be found in Vermunt and Magidson (2005).  
25 Using decade-average income smooths out fluctuations in annual incomes that could make age dependence less powerful. This possibility is best examined in the stochastic wage paths extension below (Case 3), where wages are more free to fluctuate across ages. Simulations in that model suggest that this paper’s results are robust to this concern.
The wage paths shown in Figure 1 are similar to those implied by the results of Fullerton and Rogers (1993), shown in their Table 4.11. Now I discuss the specification of the models’ parameter values.

### 1.4.2 Parameter specification

I assume the period utility function

\[ U(c, l) = \ln c - \frac{1}{\sigma} l^\sigma, \]

and set \( \sigma = 3 \), which implies a constant-consumption elasticity of labor supply of 0.5. The results described below are robust to alternative parameterizations, with the welfare gains from age dependence increasing when utility from consumption is more concave and the elasticity of labor supply is greater. I use an annual gross rate of return of five percent, which implies that \( R = (1.05)^{10} \) because I am using decade-long age ranges.

I assume constant Pareto weights \( \alpha^i = 1.00 \) for all \( i \), which implies a pure Utilitarian social welfare function. More redistributive assumptions increase the power of age-dependent taxes to raise welfare.

With these data and parameters, I simulate the policy models. Now, I turn to the results of these simulations.

### 1.4.3 Simulation results

I focus on four outputs from the numerical simulations: intratemporal distortions, average tax rates, intertemporal distortions, and welfare. For each, I compare the results under the Static Mirrlees, Partial Reform, and Full Optimum policy scenarios.

**Intratemporal distortions** First, consider intratemporal distortions as defined in (8). Table 2 lists these distortions by age and type.
Table 2: Intratemporal distortions in baseline model simulation

<table>
<thead>
<tr>
<th>Intratemporal distortion</th>
<th>Age range</th>
<th>Bottom</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Top Expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
<td>30-39</td>
<td>0.30</td>
<td>0.35</td>
<td>0.21</td>
<td>0.25</td>
<td>0.44</td>
<td>0.35</td>
<td>0.18</td>
<td>0.29</td>
<td>0.42</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>0.26</td>
<td>0.45</td>
<td>0.27</td>
<td>0.27</td>
<td>0.41</td>
<td>0.26</td>
<td>0.30</td>
<td>0.29</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>0.14</td>
<td>0.49</td>
<td>0.38</td>
<td>0.26</td>
<td>0.45</td>
<td>0.30</td>
<td>0.23</td>
<td>0.37</td>
<td>0.27</td>
<td>0.00</td>
</tr>
<tr>
<td>Partial Reform</td>
<td>30-39</td>
<td>0.22</td>
<td>0.37</td>
<td>0.17</td>
<td>0.27</td>
<td>0.36</td>
<td>0.17</td>
<td>0.21</td>
<td>0.27</td>
<td>0.17</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>0.27</td>
<td>0.41</td>
<td>0.27</td>
<td>0.27</td>
<td>0.40</td>
<td>0.21</td>
<td>0.27</td>
<td>0.34</td>
<td>0.27</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>0.31</td>
<td>0.37</td>
<td>0.46</td>
<td>0.23</td>
<td>0.37</td>
<td>0.33</td>
<td>0.20</td>
<td>0.35</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>Full Optimum</td>
<td>30-39</td>
<td>0.23</td>
<td>0.39</td>
<td>0.25</td>
<td>0.19</td>
<td>0.38</td>
<td>0.23</td>
<td>0.14</td>
<td>0.30</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>0.25</td>
<td>0.39</td>
<td>0.26</td>
<td>0.24</td>
<td>0.38</td>
<td>0.25</td>
<td>0.16</td>
<td>0.32</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>0.27</td>
<td>0.37</td>
<td>0.25</td>
<td>0.27</td>
<td>0.40</td>
<td>0.03</td>
<td>0.24</td>
<td>0.31</td>
<td>0.26</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The most striking difference between the scenarios is the treatment of the high income young (individuals in the 30-39 age range). Figure 2 plots the distortions for these workers against annual income.26 The lines in the figures connect discrete points corresponding to the I types at each age; the jagged pattern of distortions is a consequence of this discreteness.

Figure 2: Intratemporal Distortions on Young Workers

The Static Mirrlees policy substantially distorts the intratemporal choices of high-wage young workers, while the Partial Reform and Full Optimum policies do not.27 This is the numerical counterpart to Proposition 1, in which we saw that the classic result of no marginal distortion at the top of the income distribution fails to extend to the Static Mirrlees policy within age groups. To repeat the intuition, a distortion allows the planner to collect more tax revenue from higher earners. Distortions are therefore valuable on all except the top earner across ages in the Static Mirrlees policy. But, they bring no benefits for age-dependent policies when levied on the top earners in each age group, because individuals are restricted to age-specific tax schedules.

---

26 Income from the simulation results is converted to annual U.S. dollars as follows. The median annual hours worked in the data is 2,070 per year, while the Partial Reform planner has the corresponding worker exert 0.84 units of labor effort. This implies that a worker exerting one unit of labor effort per period in the model would work approximately 2,477 hours per year. I use this number as the benchmark for normal hours per year, and multiply the simulation results for income by it to obtain annual income as shown.

27 The Static Mirrlees distortions differ across age groups even though its taxes are age-independent because the intratemporal distortion \( \tau(i,t) \) depends on an individual’s wage. Two individuals of different ages with different wages who choose the same income and consumption allocation will have different implied distortions.
More generally, the use of intratemporal distortions decreases as the sophistication of policy increases. The final column of Table 2 shows that the unweighted average marginal distortion across all types and ages is 0.306 in the Static Mirrlees policy, 0.285 in the Partial Reform policy, and 0.261 in the Full Optimum policy.\footnote{If we weight distortions by income, the income-weighted average marginal distortion falls from 0.286 in the Static Mirrlees policy to 0.256 in the Partial Reform policy and 0.229 in the Full Optimum policy.} Moreover, these differences in levels are in addition to the improved pattern of distortions in the more sophisticated policies. The combination of lower average distortions and better-designed distortions encourages labor effort under the more sophisticated policies, raising total output and the efficiency of the economy. We will see the welfare impact of these efficiency gains below.

**Average taxes**  Average tax rates are substantially affected by age dependence, as well.\footnote{An individual's average tax rate is defined as the ratio $\frac{\text{Average Tax}}{\text{Income}}$.} Table 3 lists the average tax rates for each scenario.

**Table 3: Average tax rates in baseline model simulation**

<table>
<thead>
<tr>
<th>Age range</th>
<th>Bottom</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-39</td>
<td>-109.7</td>
<td>-63.9</td>
<td>-21.2</td>
<td>-17.9</td>
<td>-10.1</td>
<td>7.6</td>
<td>12.2</td>
<td>14.0</td>
<td>20.0</td>
<td>28.5</td>
</tr>
<tr>
<td>40-49</td>
<td>-109.7</td>
<td>-63.9</td>
<td>-17.9</td>
<td>-17.9</td>
<td>-0.4</td>
<td>12.2</td>
<td>14.0</td>
<td>18.8</td>
<td>28.5</td>
<td>39.6</td>
</tr>
<tr>
<td>50-59</td>
<td>-109.7</td>
<td>-63.9</td>
<td>-17.9</td>
<td>-21.2</td>
<td>-0.4</td>
<td>13.1</td>
<td>12.2</td>
<td>19.2</td>
<td>28.5</td>
<td>37.0</td>
</tr>
<tr>
<td>30-39</td>
<td>-125.1</td>
<td>-82.9</td>
<td>-30.6</td>
<td>-28.5</td>
<td>-15.5</td>
<td>4.8</td>
<td>6.5</td>
<td>9.8</td>
<td>20.9</td>
<td>27.0</td>
</tr>
<tr>
<td>40-49</td>
<td>-120.6</td>
<td>-63.7</td>
<td>-15.9</td>
<td>-15.9</td>
<td>3.3</td>
<td>18.3</td>
<td>19.5</td>
<td>22.6</td>
<td>33.9</td>
<td>42.7</td>
</tr>
<tr>
<td>50-59</td>
<td>-135.1</td>
<td>-54.1</td>
<td>-21.1</td>
<td>-21.1</td>
<td>6.1</td>
<td>17.5</td>
<td>17.2</td>
<td>26.4</td>
<td>35.8</td>
<td>46.5</td>
</tr>
<tr>
<td>30-39</td>
<td>-123.4</td>
<td>-87.0</td>
<td>-35.0</td>
<td>-18.9</td>
<td>-17.3</td>
<td>-1.0</td>
<td>9.6</td>
<td>6.9</td>
<td>16.1</td>
<td>9.7</td>
</tr>
<tr>
<td>40-49</td>
<td>-133.9</td>
<td>-70.6</td>
<td>-20.7</td>
<td>-21.4</td>
<td>2.7</td>
<td>14.9</td>
<td>25.3</td>
<td>25.2</td>
<td>37.5</td>
<td>49.9</td>
</tr>
<tr>
<td>50-59</td>
<td>-155.3</td>
<td>-62.7</td>
<td>-10.0</td>
<td>-35.7</td>
<td>4.5</td>
<td>31.4</td>
<td>7.0</td>
<td>32.0</td>
<td>41.1</td>
<td>57.9</td>
</tr>
</tbody>
</table>

Though not immediately apparent from Table 3, the average tax schedule is necessarily the same for all age groups under the Static Mirrlees policy. In the Partial Reform and Full Optimum, separate average tax schedules face workers in their thirties, forties, and fifties. In particular, Figure 3 shows that workers face lower average taxes in their thirties than later in life under the optimal age-dependent tax policy.

**Figure 3: Average Tax Rates in Partial Reform of Baseline Economy**

\[\text{Average Tax Rate} = \frac{\text{Average Tax}}{\text{Income}}\]
Figure 3 shows that the magnitude of the difference across ages in average tax schedules can be substantial. In the Partial Reform policy, a worker earning $55,000 faces an average tax rate in his thirties that is more than 7 percentage points lower than in his forties.

Why do the more sophisticated planners use lower average taxes on individuals when they are young? The data show wages rising from the thirties to the forties in all income groups. Individuals want to borrow against future wages to raise consumption when young, but in this baseline economy they cannot transfer resources across periods. Age-dependent tax policy can substitute for private borrowing by lowering average taxes when wages are low: i.e., in workers’ thirties. The Static Mirrlees planner cannot do so, because it cannot target lower average taxes at an age group.

**Intertemporal distortions** Next, Table 4 shows the intertemporal distortions under each policy. Here, the main advantage of history dependence is made clear, as only the Full Optimum policy avoids these distortions entirely, providing fully smoothed consumption to all workers.

<table>
<thead>
<tr>
<th>Age range</th>
<th>Bottom</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>30–39</td>
<td>1.00</td>
<td>1.00</td>
<td>1.04</td>
<td>1.00</td>
<td>1.08</td>
<td>1.14</td>
<td>1.07</td>
<td>1.13</td>
<td>1.30</td>
<td>1.47</td>
</tr>
<tr>
<td>40–49</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
<td>1.00</td>
<td>1.03</td>
<td>0.94</td>
<td>1.02</td>
<td>1.00</td>
<td>1.21</td>
</tr>
<tr>
<td>50–59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–39</td>
<td>0.94</td>
<td>0.96</td>
<td>0.96</td>
<td>0.94</td>
<td>0.98</td>
<td>1.00</td>
<td>1.01</td>
<td>1.02</td>
<td>1.06</td>
<td>1.26</td>
</tr>
<tr>
<td>40–49</td>
<td>0.98</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>1.02</td>
<td>1.01</td>
<td>0.94</td>
<td>1.03</td>
<td>1.01</td>
<td>1.07</td>
</tr>
<tr>
<td>50–59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–39</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>40–49</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>50–59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While not smoothing for each worker, the Partial Reform policy smooths in aggregate across ages, as shown formally in the Symmetric Inverse Euler Equation above, expression (16). This improves on the Static Mirrlees policy, which we can see in Table 4 by noting that the distortions to the intertemporal margin are smaller, on average, in the Partial Reform than in the Static Mirrlees policy.

**Welfare gain and decomposition** Finally, I quantify and identify the sources of the welfare gain from reform. For each policy, Table 5 lists overall social welfare and lifetime utility by type in consumption equivalent units. Note that the values in Table 5 are best used to see who gains and loses from more sophisticated policies, not to see how much one gains or loses: i.e., they are meaningfully ordinally, not cardinally.

<table>
<thead>
<tr>
<th>Type</th>
<th>Social Welfare (consumption equivalents*)</th>
<th>Lifetime Utility (consumption equivalents**)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
<td>2.507</td>
<td>2.33</td>
</tr>
<tr>
<td>Partial Reform</td>
<td>2.524</td>
<td>2.28</td>
</tr>
<tr>
<td>Full Optimum</td>
<td>2.530</td>
<td>2.30</td>
</tr>
</tbody>
</table>

The main findings are that age dependence generates a large welfare gain in absolute size and that it captures a large majority of the gain from full reform to a more complex, optimal dynamic policy. Moreover, age dependence yields a more equal distribution of utility than is possible under Static Mirrlees policy.

30 The simulation results show that age dependence makes consumption less smooth for low earners. The reason for this was discussed in the section on intertemporal distortions: when smoothing high-earners’ consumption, the planner skews low-earners’ consumption in the opposite direction to most efficiently satisfy incentives.
First, the increase in welfare due to age dependence alone is large, equivalent to a 1.8 percent increase in aggregate consumption or roughly $200 billion in current U.S. dollars, annually. Specifically, if the Static Mirrlees planner received a windfall enabling it to increase each individual’s consumption by 1.8 percent while holding labor effort fixed, welfare in the Static Mirrlees policy would equal that in the Partial Reform policy. Below, I provide a detailed decomposition of this large welfare gain.

Second, the gain from this Partial Reform captures 73 percent of the gain from reform to the Full Optimum. Specifically, the additional welfare gain from history dependence is 0.7 percent of aggregate consumption, so that the total gain due to reform from the Static Mirrlees to the Full Optimum is 2.5 percent of aggregate consumption. As discussed more below, the advantage of history dependence is not larger in this baseline model because many of the empirical wage paths shown in Figure 1 have similar shapes.

Finally, this welfare gain is concentrated among the low-skilled, so that the Partial Reform and Full Optimum policies achieve more egalitarian (and nearly identical) distributions of lifetime utility across income groups. Table 5 shows that the Static Mirrlees policy provides higher utility to types 6 through 10 while the more sophisticated policies increase utility among individuals in the bottom half of the income distribution.31

What drives the welfare gain from age dependence? Figure 4 shows the results of a welfare gain decomposition that attributes the gain from Partial Reform to improvements in efficiency, equity, and consumption-smoothing.

![Figure 4: Decomposition of Welfare Gains from Partial Reform](image)

Note: Components sum to 93%

Nearly half of the welfare gain from age dependence is because the economy is more efficient under age-dependent taxes. In Table 2, we saw that the average marginal distortion to labor effort is lowered by adding age dependence to a static Mirrleesian policy. This encourages more effort, so that total output is 2.4 percent higher under the Partial Reform policy than under the Static Mirrlees policy. To estimate the welfare impact of this increase, consider a simple thought experiment. Take the Static Mirrlees allocation of income and consumption and suppose that each individual were required to earn and allowed to consume 2.4 percent more at each age. The welfare gain of reform from this modified version of the Static Mirrlees

---

31 Note that the welfare gains from age dependence do not depend on funding gains for the poor with losses by the rich. In Section 5, I show that age dependence generates only slightly smaller welfare gains when constrained to be Pareto-improving.
policy to the Partial Reform policy is slightly more than 1.0 percent of aggregate consumption. Thus, the increased output due to efficiency gains accounts for approximately 43 percent of the total welfare gain from age dependence.

Most of the remaining welfare gain from age dependence is due to a more equitable distribution of resources than under an age-independent policy. In particular, the Partial Reform planner allocates consumption to individuals with higher marginal utilities of consumption and requires production from individuals with lower marginal disutilities of income than does the Static Mirrles planner. I separately estimate the welfare impacts of each of these two factors.

To estimate the effect of the distribution of consumption, consider an experiment in which each individual’s consumption path under the Static Mirrles policy is scaled to provide the same share of total consumption (in present value) as under the actual Partial Reform policy. This hypothetical allocation replicates the Partial Reform’s allocation of consumption across individuals while holding fixed the Static Mirrles level of total consumption. Specifically, it raises the present value of consumption for the lowest income group by almost 3 percent relative to the actual Static Mirrles policy, an increase offset by lower present values of consumption for higher income groups. Because utility from consumption is concave, this hypothetical Static Mirrles policy yields higher welfare than does the actual Static Mirrles policy, and the consumption-equivalent welfare gain from this hypothetical Static Mirrles to the Partial Reform is only 1.52 percent of Static Mirrles output. Thus, the distribution of consumption in accordance with the Partial Reform policy accounts for approximately 17 percent of the welfare gain from age dependence (0.3 of the 1.8 percent of total consumption-equivalent gain).

To estimate the impact of allocating required income to those with lower marginal disutilities of labor effort, I consider an analogous experiment to that for consumption. I scale the income required from each individual in the Static Mirrles to equal (in present value) the same share of total income as in the Partial Reform. Similar calculations to those for consumption imply that the distribution of required income in accordance with the Partial Reform policy accounts for approximately 27.5 percent of the welfare gain from age dependence.

Finally, age dependence allows for more efficient intertemporal allocations, i.e., more consumption-smoothing, than in the Static Mirrles. Consider an experiment in which each individual’s present value of consumption under the Static Mirrles policy is allocated across ages as it is in the actual Partial Reform policy. This hypothetical Static Mirrles policy achieves higher social welfare than the true Static Mirrles policy, implying that the Partial Reform’s increased consumption-smoothing accounts for about 5.5 percent of the gain from age dependence.

Why do these gains capture nearly three-quarters of the gain from reform to the Full Optimum? The Full Optimum’s advantage over Partial Reform is history dependence. History dependence is most valuable when wage paths cross or have substantially different slopes, as history-independent policies then have to address incentive problems that vary substantially by age. In contrast, the Full Optimum planner’s ability to track individuals allows it to target redistribution and smooth consumption despite differently-sloped wage paths. In the data used for the baseline simulation, as shown in Figure 1, most of the wage paths have similar shapes, thereby reducing the benefit from history dependence. In Section 3, I consider an extension of the baseline model that incorporates crossing wage paths, as guided by the data, by allowing wage paths to be stochastic rather than deterministic. As discussed there, the Partial Reform policy continues to capture a

---

32Note that it also violates the incentive constraints on the Static Mirrles problem, which is why the Static Mirrles planner could not achieve this hypothetical allocation even though it satisfies the feasibility constraint.
large majority of the gain from the Full Optimum in that extension.

1.5 Summary of baseline model

In this section, I theoretically and numerically characterized age-dependent taxation in a baseline model. I found that age dependence substantially improves on age-independent policy along the two margins that dominate theoretical Mirrleesian tax analysis: the intratemporal consumption-leisure margin and the intertemporal savings margin. Then, I used detailed individual wage data to show the effects of age dependence quantitatively. Age dependence generates a welfare gain estimated at 1.8 percent of aggregate annual consumption and captures 73 percent of the gains from reform to the dynamic optimum. In the next several sections, I show that these results are largely robust to extensions of this baseline model to more complicated economic environments.

2 Case 2: Model with private saving and borrowing

In this section, I examine how the results from the baseline model are affected by allowing individuals to save and borrow across ages. Private saving and borrowing generate an important new set of incentive problems for policy, in that individuals may now subsidize consumption with after-tax income earned at a different age. This affects the marginal tradeoffs facing individuals at each age and, therefore, the optimal policy toward them.

Before showing how the three policy scenarios respond to private transfers of resources across periods, it is important to clarify my assumptions on how the three policy scenarios can respond. In particular, I need to specify whether capital taxation is available to each policy and what forms it can take. While it is natural to assume that there are no restrictions on capital taxation for the Full Optimum policy, it is less clear what the appropriate assumption is for the Static Mirrlees and Partial Reform policies.

I assume that the Static Mirrlees and Partial Reform planners can neither tax nor subsidize private saving or borrowing in any way. This is a conservative assumption when gauging the power of age dependence, in that it maximizes both the potential for private saving and borrowing to undermine the baseline results and the relative power of the Full Optimum, which has unlimited flexibility in taxing and subsidizing intertemporal transfers. For example, if I allow the Partial Reform and Static Mirrlees policies to include a 15 percent tax rate on capital income (resembling the current U.S. system for capital gains and dividends), the absolute and relative sizes of the welfare gains from Partial Reform increase relative to the results below.

Thus, Partial Reform is defined consistently throughout the paper: it always means only that labor income taxes can depend on age. One interesting extension to this paper’s analysis would be to consider age-dependent linear capital taxation, which would increase the potential power of age dependence.

2.1 Analysis of the social planner’s problem in three policy scenarios

As in the baseline model, the social planner specifies a menu of bundles to maximize social welfare subject to feasibility and incentive constraints. With private saving and borrowing, these bundles are of pre-tax income and after-tax income, not consumption. For brevity, I relegate the formal statement of the planner’s

---

33 A technical note: I assume that savings and debt are observable to the planner. The term "private" indicates private sector, not "hidden," which has a specific meaning in the optimal tax literature.
problems to the Technical Appendix and discuss their key components here.34

Private saving and borrowing complicates the incentive constraints facing the planner in the two policies that lack history dependence. In the Static Mirrlees policy, individuals are free to choose any path of after-tax incomes, including those that are intended for individuals of different types and ages, and transfer them across periods by saving and borrowing. As in the baseline model, the Partial Reform planner has the advantage of conditioning taxes on age, simplifying its problem, but individuals now have the ability to smooth consumption on their own. Meanwhile, the planner’s problem for the Full Optimum scenario is unchanged from the baseline model. Because it can link allocations across an individual’s lifetime, the Full Optimum planner spreads the after-tax income received by an individual over its lifetime optimally, leaving the individual’s optimal choice undistorted.

How do these more complicated incentive constraints affect policy? In the baseline model, I analyzed policy along two margins: the intratemporal and intertemporal. In this model, with private saving and borrowing, the second of these margins goes undistorted in all three policy scenarios.35 Therefore, I focus my analysis in this section on the intratemporal margin.

The impact of private saving and borrowing on intratemporal distortions can best be seen in a simple example. Consider an economy with only two worker types, \( i = \{L, H\} \) for low and high skilled, and two ages \( t = \{1, 2\} \). Suppose that the high-skilled type \( H \) is always higher-skilled than the low-skilled type \( L \), so that \( w_{1H}^t > w_{1L}^t \) and \( w_{2H}^t > w_{2L}^t \). Finally, assume that the utility function takes the simple form used in the numerical simulations of the baseline model:

\[
U(c, y) = \ln c - \frac{1}{\sigma} \left( \frac{y}{w} \right)^\sigma.
\]

The Full Optimum policy’s treatment of the high-skilled worker is unchanged by private saving and borrowing. Thus, we know from Proposition 1 (Top Marginal Distortion) that the high-skilled worker in this example will face a zero intratemporal distortion at both ages in the Full Optimum policy.

The Partial Reform policy, however, may respond to the new incentive problems by distorting the high-skilled worker. Formal and numerical examples of this are provided in the Technical Appendix. Intuitively, the high-skilled worker is tempted to save some of its after-tax income from the first period and use that savings to raise its consumption while working less and claiming the tax treatment of the low-skilled worker later in life. If the Partial Reform policy distorts the young, high-skilled’s intertemporal margin, it discourages him or her from earning the extra income that funds oversaving. The logic works in the other direction as well. If old, high-skilled workers are tempted to overborrow and work less when young, the planner can make such cheating more costly by distorting the high-skilled worker’s intratemporal margin when old.

There is an enlightening relationship between the Partial Reform planner’s use of intratemporal distortions and the well-known Inverse Euler Equation in stochastic dynamic optimal tax models, first noted in Rogerson (1985) and prominently explored by Golosov, Kocherlakota, and Tsyvinski (2003). The Inverse Euler Equation suggests that policy ought to distort the after-tax return to saving (ex post, as shown in Kocherlakota, 2005) in order to counteract the temptation people face in the presence of skill shocks to oversave and falsely claim a low skill level, and thus lower taxes, in the future. The intratemporal distortion

---

34 Also in the Technical Appendix, I show that the allocations derived below using a direct mechanism are implementable through a nonlinear tax system, as in the Baseline case. I also show that the intratemporal distortions can be mapped to equivalent marginal taxes on earned income. The key to this result is that individuals smooth any extra after-tax earned income, so the utility value of a dollar of after-tax income is equal to the marginal utility of consumption for the individual, and the standard expression for the intratemporal distortion is thus equal to the marginal tax rate.

35 The Static Mirrlees and Partial Reform policies cannot distort the intertemporal margin by assumption, as discussed above. The Full Optimum chooses not to, as proven in Section 1.
above works toward the same goal, even though wages are fully deterministic in this economy. The Partial Reform planner, lacking history dependence, must act as if wages are stochastic even when they are not. Though, by assumption, it has no way to tax savings in the Partial Reform, it uses intratemporal distortions to try to achieve the same results.

The use of intratemporal distortions to substitute for intertemporal distortions puts into jeopardy, in principle, the result from the baseline model that age dependence lowers marginal distortions on high wage young workers. To check whether that result holds in Case 2, and to test the robustness of the other main lessons from the baseline model, I turn to numerical simulations.

### 2.2 Numerical results

For the numerical simulation of this model, I use the same approach as in the baseline economy simulation. Computational considerations cause me to reduce the number of types of individuals to \( I = 7 \). The same underlying data and latent class approach are used as in the baseline case. Table 6 shows the wage paths for the seven types along with descriptive demographic data.

| Table 6: Data for simulation of Case 2: private saving and borrowing |
|-------------------------|---|---|---|---|---|---|---|
| **Wage paths** | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **Average wage by age range ($1999)** | | | | | | | |
| 30-39 | 7.31 | 10.84 | 15.02 | 15.90 | 20.36 | 27.44 | 36.81 |
| 40-49 | 7.31 | 11.47 | 16.87 | 17.19 | 23.03 | 33.02 | 49.89 |
| 50-59 | 7.03 | 11.57 | 17.27 | 15.74 | 24.16 | 35.14 | 53.05 |

**Descriptive data**
- Proportion of population: 0.23, 0.18, 0.14, 0.16, 0.15, 0.09, 0.05
- Proportion race=white: 0.41, 0.52, 0.64, 0.73, 0.76, 0.86, 0.93
- Proportion with college degree: 0.05, 0.05, 0.12, 0.25, 0.24, 0.52, 0.70
- Proportion gender=male: 0.59, 0.71, 0.82, 0.90, 0.90, 0.94, 0.95

The intratemporal distortions in each policy are shown in Table 7.

| Table 7: Intratemporal distortions in Case 2 model simulation |
|---------------------|---|---|---|---|---|---|---|---|---|---|
| **Intratemporal distortion** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | **Expected value** |
| **Static Mirrlees** | | | | | | | | | 0.285 |
| 30-39 | 0.24 | 0.30 | 0.22 | 0.24 | 0.38 | 0.41 | 0.17 |
| 40-49 | 0.24 | 0.41 | 0.31 | 0.34 | 0.23 | 0.18 | 0.20 |
| 50-59 | 0.14 | 0.43 | 0.36 | 0.21 | 0.34 | 0.32 | 0.22 |
| **Partial Reform** | | | | | | | | | 0.277 |
| 30-39 | 0.19 | 0.31 | 0.16 | 0.27 | 0.28 | 0.27 | 0.25 |
| 40-49 | 0.29 | 0.39 | 0.22 | 0.26 | 0.31 | 0.28 | 0.02 |
| 50-59 | 0.32 | 0.36 | 0.43 | 0.24 | 0.30 | 0.25 | 0.00 |
| **Full Optimum** | | | | | | | | | 0.241 |
| 30-39 | 0.24 | 0.33 | 0.24 | 0.18 | 0.26 | 0.20 | 0.00 |
| 40-49 | 0.25 | 0.35 | 0.22 | 0.19 | 0.28 | 0.24 | 0.00 |
| 50-59 | 0.26 | 0.35 | 0.14 | 0.23 | 0.28 | 0.24 | 0.00 |

Figure 5 plots these for individuals in their thirties, as in the baseline analysis.
As in the baseline model, most high-income young continue to face larger distortions to labor supply in the age-independent Static Mirrlees than in the more sophisticated policies. The highest-earning young, however, face a slightly larger distortion in the Partial Reform than in the Static Mirrlees. As discussed above, this reflects the Partial Reform policy’s use of intratemporal distortions to discourage the deviation strategy in which individuals oversave when young, planning to use the return to savings to supplement consumption while working less and claiming a more generous tax treatment when older.

Also as in the baseline model, the use of intratemporal distortions in general decreases as the sophistication of policy increases. The final column of Table 7 shows that the unweighted average marginal distortion is 0.285 in the Static Mirrlees policy, 0.277 in the Partial Reform policy, and 0.241 in the Full Optimum policy. As in the baseline, the combination of lower average distortions and a better-designed pattern of distortions encourages labor effort under the more sophisticated policies, raising total output and the efficiency of the economy.

Average taxes in this model are, in contrast to the baseline model, indeterminate in the Partial Reform policy. As individuals can freely transfer resources across ages, any pattern of average taxes can be replaced with another that transfers resources lump-sum from one age to another without affecting any individual’s choices, aggregate welfare, the allocation’s feasibility, or incentive constraints. Thus, adding private saving and borrowing does not contradict the recommendation from the baseline case to have the young face a lower average tax schedule.

Finally, the Partial Reform continues to capture a large absolute welfare gain and a substantial share of the potential gains from more comprehensive reform. Table 8 shows social welfare and each income quintile’s lifetime utility under the three policies.

---

36 The difference is greater when we weight distortions by income: the income-weighted average marginal distortion falls from 0.279 in the Static Mirrlees policy to 0.232 in the Partial Reform policy and 0.169 in the Full Optimum policy.

37 I am grateful to Ivan Werning for suggesting this result.
Reform from the Static Mirrlees policy to the Partial Reform policy yields a gain of 1.4 percent of aggregate consumption, slightly less than in the baseline model. Because the Full Optimum planner is better able to respond to the new incentive problems introduced by private saving and borrowing, this gain makes up somewhat less of the gain from full reform. Nevertheless, it captures a substantial share, 57 percent, of that potential gain. More sophisticated capital taxation would magnify the power of age dependence. For instance, in an extension with a uniform 15 percent tax on capital income (not shown), the welfare gain from Partial Reform rises to 1.6 percent of aggregate consumption, comprising 62 percent of the potential gain from the Full Optimum.

As in the baseline model, the Partial Reform’s higher overall social welfare is also shared more equally among the individuals in the population. Table 8 shows that the Partial Reform policy produces a more egalitarian distribution of utility than does the Static Mirrlees, though not as egalitarian as the Full Optimum.

### Table 8: Welfare in Case 2 model simulation

<table>
<thead>
<tr>
<th>Type</th>
<th>Social Welfare (consumption equivalents*)</th>
<th>Lifetime Utility (consumption equivalents**)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
<td>2.86</td>
<td>2.26 2.36 2.51 2.52 2.69 2.91 3.20</td>
</tr>
<tr>
<td>Partial Reform</td>
<td>2.89</td>
<td>2.36 2.46 2.57 2.58 2.72 2.91 3.19</td>
</tr>
<tr>
<td>Full Optimum</td>
<td>2.90</td>
<td>2.35 2.42 2.52 2.52 2.67 2.86 3.12</td>
</tr>
</tbody>
</table>

* The value for consumption that, if provided freely to all workers at each age, would generate the same social welfare as the actual allocation.

** The value for consumption that, if provided freely to the worker at each age, would generate the same lifetime utility as the worker obtains with the actual allocation.

3 Case 3: Model with stochastic wage paths

In this section, I return to a setting in which individuals cannot transfer resources across time, but I explore a new variation on the baseline model by modeling wage paths as stochastic rather than deterministic. Stochastic wage paths generalize the baseline model in two important ways. First, they mean that individuals and the planner are uncertain about their future wages. As shown in Rogerson (1985) and Golosov, Kocherlakota, and Tsyvinski (2003), this uncertainty affects individuals’ labor supply and saving and borrowing behavior, with important implications for dynamic optimal policy. Second, they allow for substantially more heterogeneity in wage paths. In the baseline model, the representative wage paths had generally similar shapes. In this section and the next, individuals are not classified into types, so wage paths will have a variety of shapes.

The extent to which individuals’ wage paths are determined over time due to stochastic shocks, rather than at the start of their working lives, is the subject of substantial recent research. Keane and Wolpin (1997) and Storesletten, Telmer, and Yaron (2001), provide evidence that stochastic shocks account for as little as 10 percent and as much as 40 percent of total variation in wage paths, respectively. Guvenen (2007) finds evidence that individuals undergo substantial learning over time about the shape of their wage path. The larger the role of uncertainty in wage paths, the more important it is to understand how stochasticity affects the power of age dependence.

I model stochastic wages as a simple Markov process. At each age of working life, individuals are distributed among an age-specific set of discrete wage levels. A separate transition matrix links each age’s wage distribution to the next, so that a transition matrix between ages $t$ and $t+1$ determines the distribution of all individuals with a given wage at age $t$ among the set of wage levels at age $t+1$. This simple Markov
approach yields a transparent and computationally tractable representation of the dynamic uncertainty and heterogeneity in wage paths from the data.

As with the baseline model, I begin with a theoretical analysis of the policy scenarios as social planners’ problems. The key result is Proposition 3 \((\text{Baseline and Case 3 Equivalence})\), which shows that the Static Mirrlees and Partial Reform planners’ problems in this Case 3 model are identical to their problems in an appropriately-specified baseline model from the baseline case. Then, I use numerical simulations to show that the quantitative results from the baseline model carry through to this model with stochastic wage paths.

### 3.1 Analysis of the social planner’s problem in three policy scenarios

As in the baseline model, I work with social planners’ problems in three policy scenarios: the Static Mirrlees, Partial Reform, and Full Optimum.

I begin by defining some notation. Denote an individual’s true path of wages as \(W^{(t)}_T = \{ w^{(t)}_1, w^{(t)}_2, \ldots w^{(t)}_T \} \) and let \(\pi^{(t)}_i\) denote the population proportion represented by this individual. Using these probabilities, let \(\pi^{(t)}_i = \sum_{j=1}^T p^{(t)}_{i,j} \pi^{(t)}_j\) denote the probability of wage level \(w^{(t)}_i\) at age \(t\). It is the sum of the population proportions of the individuals whose wage paths equal \(w^{(t)}_i\) at age \(t\). Note that \(\sum_{j=1}^T \pi^{(t)}_j = 1\) for all \(t\). Denote the transition matrix between ages \(t\) and \(t+1\) as \(P_{t,t+1}\), whose element \((m,n)\) is:

\[
P_{t,t+1}(m,n) = \Pr(w_{t+1}^m | w_t^n).
\]

In words, \(P_{t,t+1}(m,n)\) is the probability that an individual with wage \(w_t^n\) will have wage \(w_{t+1}^m\). Thus, the population proportion of an individual with wage path \(W^{(t)}_T\) can also be written \(\pi^{(t)}_i = \prod_{t=1}^{T-1} \pi^{(t)}_i P_{t,t+1} (i_t, i_{t+1})\).

The structure of each planner’s problem is the same as in the previous sections. To maximize social welfare, each planner offers a menu of income and consumption pairs to individuals. In the Static Mirrlees and Partial Reform policies, the planner offers a pair \(\{c^i_t, y^i_t\}\) as the consumption and income intended to be chosen by an individual with wage \(w^i_t\) at age \(t\). In the Full Optimum policy, the allocations can be history-dependent.\(^{38}\) Social welfare depends on the Pareto weights \(\alpha^j\) assigned to individuals with different wage paths. In the numerical simulations below, I assume that weights are constant across all individuals, so that the planner is a pure Utilitarian, but more complicated assumptions can be incorporated.\(^ {39}\)

The formal statement of the Static Mirrlees and Partial Reform planner’s problems can be found in the Technical Appendix. The following proposition gives the key result regarding them.

**Proposition 3 (Baseline and Case 3 Equivalence)** Consider a stochastic wage path \(W^{(t)}_T\) and a deterministic wage path \(W^j_T\). If, for each \(i(t)\), there exists a \(j\) such that \(\pi^{(t)}_i = \pi^j\) and \(w^{(t)}_i = w^j_t\) for all \(t\), where \(w^{(t)}_i \in W^{(t)}_T\) and \(w^j_t \in W^j_T\), then:

- the solution to the Static Mirrlees planner’s problem in the baseline model is also the solution to the Case 3 Static Mirrlees planner’s problem,

\(^{38}\)The Static Mirrlees and Partial Reform allocations may, in principle, be different for two individuals with different histories but the same current wage, in that these individuals could choose different \((c, y)\) pairs. In this Case 3 model, however, individual decisions depend only on their current wage, so this possibility is irrelevant.

\(^{39}\)See the Technical Appendix for a discussion of this.
the solution to the Partial Reform planner’s problem in the baseline model is also the solution to the
Case 3 Partial Reform planner’s problem.

Proof. In Technical Appendix. ■

This proposition considers a deterministic economy that has the same set of individual wage paths as
the stochastic economy will have, ex post. Its implications hold because, in each of these policy scenar-
ios, the objective function, feasibility constraint, and incentive constraints are the same with stochastic or
deterministic wage paths. Thus, the same distributions of $c$ and $y$ solve the planner’s problems in each
model.

The key to this result is that adding stochasticity in wages fails to change the problem for either the
planner or the individuals in the Static Mirrlees and Partial Reform scenarios. For these planners, being
restricted to history-independence plays the same role as wage stochasticity. When a planner cannot track
individuals across ages, it must satisfy incentives age-by-age, rather than over a lifetime, so it is already
setting policy as if wages were stochastic. For individuals, the addition of stochasticity has no e¤ect on their
incentives relative to the baseline model because, in the Case 3 model, individuals cannot transfer resources
between periods and utility is separable across periods. Each age involves an isolated optimization for the
individual, so the stochasticity of wages is irrelevant.

In contrast, the Full Optimum planner’s problem is substantially a¤ected by stochasticity. Stochastic
wages multiply the number of possible wage paths, each of which is assigned a history-dependent allocation at
each age. In particular, let $\{c^i_t(W_{t-1}^{j(t)}), y_t(W_{t-1}^{j(t)})\}$ be the allocation of consumption and pre-tax income in-

tended for an individual of age $t$ who has reported the (possibly false) wage path $W_{t-1}^{j(t)} = \{w_{t-1}^{j1}, w_{t-1}^{j2}, ...w_{t-1}^{jz}\}$
and who reports the current wage $w_t^i$. The Full Optimum planner’s incentive constraints must guarantee
that individuals would rather reveal their true wage path $W_T^{i(t)}$, age by age, rather than any other path, taking into account that individuals know the true transition matrices.

The Full Optimum planner in the Case 3 model solves the following problem:

Problem 4  (Case 3 Full Optimum: Age-Dependent and History-Dependent)

$$\max_{\{c,y\}} \left\{ \sum_{i(t)} \pi^{i(t)} \alpha(W_T^{i(t)}) \sum_t \beta^{t-1} \left( u\left(c^i_t\left(W_T^{i(t)}\right)\right) - v\left(y_t\left(W_{t-1}^{i(t)}\right)\right) - \frac{y_t\left(W_{t-1}^{i(t)}\right)}{W_T^{i(t)}} \right) \right\}$$

subject to feasibility

$$\sum_{i(t)} \pi^{i(t)} \sum_{t=1}^{T} R^{T-t} \left( y_t^i\left(W_T^{i(t)}\right) - c^i_t\left(W_T^{i(t)}\right) \right) = 0.$$
Next, for all $i, j$ and $t < T$:

\[
U \left( W_{t-1}^j, w_t^i \right) + \beta \sum_{i_{t+1}} P_{t,t+1} (i, i_{t+1}) U \left( W_{t-1}^j, w_t^i, w_{t+1}^{i_{t+1}} \right) \\
\geq U \left( W_{t-1}^j, w_t^j \right) + \beta \sum_{i_{t+1}} P_{t,t+1} (i, i_{t+1}) U \left( W_{t-1}^j, w_t^j, w_{t+1}^{i_{t+1}} \right)
\]

where $U \left( W_{t-1}^j, w_t^i \right)$ is the period utility at age $t$ of an individual reporting the sequence of wages defined by \( (W_{t-1}^j, w_t^i) \), so

\[
U \left( W_{t-1}^j, w_t^i \right) = u \left( c_t^i \left( W_{t-1}^j \right) \right) - v \left( \frac{y_t^i \left( W_{t-1}^j \right)}{w_t^i} \right)
\]

\[
U \left( W_{t-1}^j, w_t^i, w_{t+1}^{i_{t+1}} \right) = u \left( c_{t+1}^{i_{t+1}} \left( W_{t-1}^j, w_t^i \right) \right) - v \left( \frac{y_{t+1}^{i_{t+1}} \left( W_{t-1}^j, w_t^i \right)}{w_{t+1}^{i_{t+1}}} \right)
\]

This planner’s problem is the stochastic analogue of the Full Optimum problems in the baseline and Case 2 models. Intuitively, the incentive constraints have two components. First, they ensure that, no matter the previous path of wage claims, individuals want to reveal their true wage in the last period of working life, age $T$. Second, they ensure that truth-telling is optimal at each age $t$ prior to the final working period, no matter the previous path of claims, given that truth-telling is optimal at the next age $t + 1$. These two steps guarantee that truthful telling is optimal at all ages for all individuals.

Applying Proposition 3 (Baseline and Case 3 Equivalence), we know that the Static Mirrlees and Partial Reform policies are straightforward analogues of those from the baseline model. In particular, the Partial Reform policy has two main advantages over the Static Mirrlees policy. First, it can tailor intratemporal distortions to the wage distribution at each age\(^{40}\); second, it can satisfy the Symmetric Inverse Euler Equation, providing a more efficient intertemporal allocation of consumption.

The Full Optimum policy differs from the baseline case in one important way: it distorts the intertemporal margin. Echoing the well-known result shown by Rogerson (1985) and Golosov, Kocherlakota, and Tsyvinski (2003), the Full Optimum allocation is described by an Inverse Euler Equation. For individual $i$ of age $t$, this expression is:

\[
\frac{1}{u' \left( c_t^i \left( W_{t-1}^j \right) \right)} = \sum_{j=1}^{I} P_{t,t+1} (i, j) \frac{P_{t,t+1} (i, j)}{u' \left( c_{t+1}^j \left( W_t^i \right) \right)}
\]

(17)

The Inverse Euler Equation sets the consumption cost to the planner of providing marginal utility for an individual at age $t$ equal to the expected consumption cost of providing marginal utility for the same individual at age $t + 1$. As before, the Partial Reform policy fails to achieve this optimal intertemporal allocation, though it improves on the Static Mirrlees by satisfying the Symmetric Inverse Euler Equation.

Now, I turn to numerical simulations to test the robustness of the baseline model’s quantitative results.

\(^{40}\)Formal expressions for these are found in the Technical Appendix. They resemble, with adjustments for stochasticity, the expressions from the Baseline scenario.
3.2 Numerical results

In this section, I simulate the Case 3 planners’ problems. First, I discuss the construction of the required data and the parameter specification. Then, I describe the results of the simulations.

3.2.1 Data and Parameters

Wages are modeled as a Markov process. Individuals are distributed among age-specific sets of discrete wage levels and move between these wage levels over time according to transition matrices linking each age to the next. Computational considerations cause me to limit the number of wage levels at each age to $I = 4$. I continue to use $T = 3$ to represent the three decades of working life.

The simulations require a wage distribution for each age and transition matrices between ages. For the wage distributions, I use the data on household heads from the PSID core sample as described in Section 1, though I restrict the sample to individuals who were observed at least twice in each age range (i.e., 30-39, 40-49, and 50-59 years of age). This leaves approximately 81,000 observations representing almost 8,000 individuals. With this sample I calculate the 5th, 35th, 65th, and 95th percentile wages within each age range and use these as the four wage levels among which individuals stochastically move. These wage levels are shown in Table 10.

Transition matrices are calculated from the data as follows. I assign percentile rankings to individual wages at each age, and for each age range (i.e., 30-39, 40-49, 50-59), I average each individual’s percentile rank for the observed years. Within each age range, I then sort individuals according to this average rank and group them into wage quartiles. This assigns each individual to a wage quartile in each age range, allowing me to calculate empirical transition probabilities that populate the transition matrices in Table 10.

Table 9: Data for simulation of Case 3 (stochastic wage paths)

<table>
<thead>
<tr>
<th>Wage levels</th>
<th>Age range</th>
<th>5th</th>
<th>35th</th>
<th>65th</th>
<th>95th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage for specified percentile and age range ($1999)</td>
<td>30-39</td>
<td>5.52</td>
<td>12.43</td>
<td>19.35</td>
<td>35.76</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>5.69</td>
<td>13.46</td>
<td>21.81</td>
<td>43.59</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>5.08</td>
<td>12.30</td>
<td>20.73</td>
<td>45.00</td>
</tr>
<tr>
<td>Initial pbb</td>
<td>0.25</td>
<td>0.30</td>
<td>0.27</td>
<td>0.18</td>
<td></td>
</tr>
</tbody>
</table>

The simulations require a wage distribution for each age and transition matrices between ages. For the wage distributions, I use the data on household heads from the PSID core sample as described in Section 1, though I restrict the sample to individuals who were observed at least twice in each age range (i.e., 30-39, 40-49, and 50-59 years of age). This leaves approximately 81,000 observations representing almost 8,000 individuals. With this sample I calculate the 5th, 35th, 65th, and 95th percentile wages within each age range and use these as the four wage levels among which individuals stochastically move. These wage levels are shown in Table 10.

Transition matrices are calculated from the data as follows. I assign percentile rankings to individual wages at each age, and for each age range (i.e., 30-39, 40-49, 50-59), I average each individual’s percentile rank for the observed years. Within each age range, I then sort individuals according to this average rank and group them into wage quartiles. This assigns each individual to a wage quartile in each age range, allowing me to calculate empirical transition probabilities that populate the transition matrices in Table 10.

While the most common movement is no movement across quartiles, about 40 percent of moderate wage earners and between 20 and 30 percent of low and high wage earners switch quartiles in each transition.

I use the same parameters as specified in Section 1, including uniform Pareto weights that reflect a pure Utilitarian social welfare function.

3.2.2 Simulation Results

The simulation results for the Case 3 planner’s problems reinforce the lessons of the baseline and Case 2 simulations. As in the baseline model, I examine intratemporal distortions, average tax rates, intertemporal
First, consider intratemporal distortions. The Static Mirrlees and Partial Reform distortions are shown in Table 11.

Table 10: Intratemporal distortions in Case 3 model simulation

<table>
<thead>
<tr>
<th>Intratemporal distortion</th>
<th>Age range</th>
<th>2</th>
<th>3</th>
<th>Top</th>
<th>Expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
<td>30-39</td>
<td>0.32</td>
<td>0.26</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>0.38</td>
<td>0.42</td>
<td>0.46</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>0.13</td>
<td>0.24</td>
<td>0.37</td>
<td>0.00</td>
</tr>
<tr>
<td>Partial Reform</td>
<td>30-39</td>
<td>0.26</td>
<td>0.25</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>40-49</td>
<td>0.31</td>
<td>0.31</td>
<td>0.32</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>0.33</td>
<td>0.33</td>
<td>0.36</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 6 plots these distortions for individuals in their thirties under the Static Mirrlees and Partial Reform policies.\[41\]

The striking disparity in the treatment of the high-income young that we saw in previous models is apparent here as well, so that high-skilled young workers are inefficiently discouraged from working by an age-independent tax system.

Moreover, the Partial Reform policy again uses marginal distortions less overall than does the Static Mirrlees policy. The expected distortion is 0.267 in the Static Mirrlees, compared to 0.238 in the Partial Reform policy. The difference in the expected distortion increases when weighted by income, from 0.192 in the Static Mirrlees to 0.129 under Partial Reform.

The results on average tax rates also resemble those from the baseline model. The average tax rates for the Static Mirrlees and Partial Reform results are given in Table 11.

\[41\] To avoid confusion, the Full Optimum distortions are not shown in the table because they are not readily comparable to the other two scenarios, as they are not functions of current income only.
Average tax rates under the Partial Reform policy are plotted against lifetime income in Figure 7.

As in the baseline model, the Partial Reform policy lowers average tax rates on workers in their thirties. The size of the gap in rates in the middle of the income distribution between the young and peak earners resembles that in the baseline case, and the intuition is the same. In Case 3, individuals cannot borrow against their higher expected future wages, so tax policy can substitute for private borrowing.

I also compare the intertemporal distortions under each scenario. Table 12 shows the ratio

\[
\frac{c^i_t}{\sum_{j=1}^{J} P_{t+1} (i,j) c^j_{t+1} (W^j_t)},
\]

which rearranges the Inverse Euler Equation in expression (17) when utility of consumption is logarithmic, for the Partial Reform and Static Mirrlees policies.

### Table 11: Average Tax Rates in Case 3 model simulation

<table>
<thead>
<tr>
<th>Age range</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>-273.5</td>
<td>-273.5</td>
<td>-273.5</td>
</tr>
<tr>
<td>2</td>
<td>-38.5</td>
<td>-38.5</td>
<td>-38.5</td>
</tr>
<tr>
<td>3</td>
<td>3.1</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Top</td>
<td>34.7</td>
<td>34.7</td>
<td>33.7</td>
</tr>
</tbody>
</table>

#### Wage quartile in each age range

<table>
<thead>
<tr>
<th>Wage quartile in each age range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
</tr>
<tr>
<td>Partial Reform</td>
</tr>
</tbody>
</table>

### Table 12: Intertemporal distortions in Case 3 model simulation

<table>
<thead>
<tr>
<th>Age range</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>0.93</td>
<td>0.92</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>0.95</td>
<td>0.94</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>0.90</td>
<td>0.84</td>
<td>...</td>
</tr>
<tr>
<td>Top</td>
<td>0.85</td>
<td>1.04</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wage quartile in each age range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
</tr>
<tr>
<td>Partial Reform</td>
</tr>
</tbody>
</table>
Though not shown in the table, this ratio is equal to one in the Full Optimum for each individual (aside from some numerical noise in the extremely unlikely paths). Apparent from the table is that the deviations of this ratio from one are larger for the Static Mirrlees planner than for the Partial Reform planner. Note, in particular, the substantial distortions on the high-skilled young by the Static Mirrlees planner.

Finally, Partial Reform continues to capture a large absolute and relative welfare gain. Table 13 shows overall social welfare and lifetime utility for individuals with four representative wage paths under the three policies.

<table>
<thead>
<tr>
<th>Wage paths</th>
<th>Always lowest wage</th>
<th>Always highest wage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Mirrlees</td>
<td>2.646</td>
<td>2.47</td>
</tr>
<tr>
<td>Partial Reform</td>
<td>2.671</td>
<td>2.53</td>
</tr>
<tr>
<td>Full Optimum</td>
<td>2.672</td>
<td>2.53</td>
</tr>
</tbody>
</table>

* The value for consumption that, if provided freely to all workers at each age, would generate the same social welfare as the actual allocation.

** The value for consumption that, if provided freely to the worker at each age, would generate the same lifetime utility as the worker obtains with the actual allocation.

Age dependence yields a welfare gain equivalent to a 2.5 percent increase in aggregate consumption relative to the Static Mirrlees policy in Case 3. Moreover, this gain captures 95 percent of the welfare gain from reform to the Full Optimum policy. The utility gains are especially substantial for the lower-income workers in this model. In Table 13, I show the lifetime utilities of four individuals with wage paths that stay in the same quartile for all three periods. As in previous models, age dependence makes the distribution of lifetime utility more equal while raising overall welfare.

Why do stochastic wage paths strengthen the relative power of Partial Reform? The result relies on two factors. First, the stochastic nature of these heterogeneous wage paths weakens the Full Optimum policy, because it now has to satisfy incentives repeatedly rather than only once. This narrows the gap between the Full Optimum and the history-independent policies that were already forced to satisfy incentives at each age. Second, the empirical magnitude of wage path heterogeneity (i.e., wage path crossing) is not large enough to change the main results. At each age, current income is a powerful enough predictor of lifetime income-earning ability that the Partial Reform policy can still redistribute on a lifetime basis by redistributing within ages.42

Thus, the results from the baseline model are robust to the inclusion of wage stochasticity, at least to the extent implied by the data used in these analyses and for a parsimonious specification of the stochastic process. In the next section, I again relax the assumption that individuals cannot save or borrow, but now in the context of stochastic wages.

42 This result does not rely on the use of incomes averaged over decade-long age ranges. Though these averages smooth incomes, the correlations between income at each age and lifetime income-earning ability in these data are nearly as large as those between the decade-long average incomes and lifetime income-earning ability. Moreover, I have simulated a version of the Case 3 model in which I use five-year age ranges. The welfare gain from Partial Reform in fact increases relative to the main results, as the wider differences between age groups raise the value of age-dependent marginal distortions and transfers across age groups.
4 Case 4: Model with stochastic wage paths and private saving and borrowing

In this section, I combine the variations on the baseline model examined separately in the previous two sections and consider a model with both stochastic wages and private saving and borrowing. I make the same modifications to the model in Case 3 as I did to the baseline model when specifying the model in Case 2. In particular, I retain the assumption from Case 2 that the Static Mirrlees and Partial Reform planners cannot tax intertemporal transfers by individuals.

As in the previous models, I consider a social planning problem for each policy, where a planner maximizes social welfare subject to feasibility and incentive compatibility constraints. When individuals can transfer resources across periods, the planners in the Static Mirrlees and Partial Reform scenarios do not control consumption directly. Thus, these planners specify pre-tax income and after-tax income bundles in Case 4, just as in Case 2. The objective function for these two policies is

$$\max_{\{x,y\}} \left\{ \sum_{i(t)} \pi^{i(t)} \alpha \left( W^{i(t)}_T \right) \sum_{t=1}^{T} \beta^{t-1} \left( u \left( c^{i(t)}_t \right) - v \left( \frac{y^{i(t)}_t}{w^{i(t)}_t} \right) \right) \right\},$$

and the feasibility constraint is:

$$\sum_{i(t)} \pi^{i(t)} \sum_{t=1}^{T} R^{T-t} \left( y^{i(t)}_t - x^{i(t)}_t \right) = 0.$$ 

Both of these expressions include, as in Case 2, after-tax income $x$, and all other notation is as before.\(^ {43}\)

As usual, variations in the incentive constraints allow us to distinguish between policy scenarios. The combination of private access to capital markets and stochasticity makes these incentive constraints quite complicated, however. Therefore, I relegate them to the Technical Appendix.

In words, the incentive constraints for the Static Mirrlees and Partial Reform scenarios reflect that an individual can choose a separate deviation strategy, including saving and borrowing, for each possible true path of wages. So, the Static Mirrlees incentive constraints must ensure that each individual prefers its allocation to any of the other allocation streams it might claim. If there are $T$ periods and $I$ wage levels, the number of these other streams is $\left( (IT)^{1+I(T-1)} - 1 \right)$. That is, in the first period, an individual can claim any of $IT$ wages, including his own. When planning for the second period, for each of the $I$ possible second period wages, he can claim any of the $IT$ wages again. The Partial Reform incentive constraints are, as usual, a subset of the Static Mirrlees scenario’s because the planner can make age-dependent allocations. Each individual in the Partial Reform must be prevented from claiming allocation streams other than her own that number only $\left[ I^{1+I(T-1)} - 1 \right]$, as she can claim $I$, not $IT$, wages for each wage level she receives.

In contrast, the Full Optimum planner’s problem is unchanged from the problem without private saving, just as its Case 2 problem was unchanged from its baseline problem. Thus, the Full Optimum planner’s problem in Case 4 is identical to the Case 3 Full Optimum planner’s problem.

The complexity of the Static Mirrlees and Partial Reform problems makes it most convenient to study

\(^{43}\) The Static Mirrlees and Partial Reform allocations may, in principle, be different for two individuals with different histories but the same current wage, in that these individuals could choose different $(x, y)$ pairs. Computational considerations prevent me from allowing for this, however, and instead I restrict these policies to allocations that are identical across two such individuals. The impact of this restriction is likely to be minimal, as economic efficiency and incentive constraints require allocations to these individuals to be similar. Moreover, this restriction has no effect on the Full Optimum policy and primarily handicaps the Partial Reform policy, causing me to, if anything, underestimate the relative gain from age dependence.
their solutions numerically, so I turn to the quantitative simulations now.

4.1 Numerical results

The computational demands of the Case 4 model are substantial, so I further limit the size of the economy. I condense the lifecycle into two age periods \((T = 2)\) covering the same range as before, so that \(t = 1\) for ages 30-44 and \(t = 2\) for ages 45-59. I also limit the number of wage levels at each age to three \((I = 3)\).

The construction of the data is the same as in Case 3, though I now choose the 5th, 50th, and 95th percentiles for each age range as the representative wage levels and classify individuals into three wage quantiles. Table 14 gives the wage levels at each age and the transition matrix between the two age ranges.

Table 14: Data for simulation of Case 4

<table>
<thead>
<tr>
<th>Wage levels</th>
<th>30-44</th>
<th>45-59</th>
<th>Initial pith</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage for specified</td>
<td>5.54</td>
<td>5.29</td>
<td>0.35</td>
</tr>
<tr>
<td>5th</td>
<td>16.00</td>
<td>16.51</td>
<td>0.39</td>
</tr>
<tr>
<td>95th</td>
<td>37.97</td>
<td>44.50</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The parameterization of the models is the same as in the previous cases. I provide results on optimal intratemporal distortions and the welfare implications of reform. As in Case 2, optimal average tax rates are indeterminate in this setting for the Static Mirrlees and Partial Reform policies.

First, consider intratemporal distortions. The distortions in the three policies are listed in Table 15.

Table 15: Intratemporal distortions in Case 4 model simulation

<table>
<thead>
<tr>
<th>Wage quantile in 30-44 age range</th>
<th>Wage quantile in 45-59 age range</th>
<th>Expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>Bottom</td>
<td>0.30 0.25 0.13</td>
</tr>
<tr>
<td>Middle</td>
<td>Middle</td>
<td>0.24 0.36 0.12</td>
</tr>
<tr>
<td>Top</td>
<td>Top</td>
<td>0.26 0.37 0.13</td>
</tr>
</tbody>
</table>

Static Mirrlees

Distortion in first age range

Table 15: Intratemporal distortions in Case 4 model simulation

<table>
<thead>
<tr>
<th>Wage quantile in 30-44 age range</th>
<th>Wage quantile in 45-59 age range</th>
<th>Expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>Bottom</td>
<td>0.30 0.25 0.13</td>
</tr>
<tr>
<td>Middle</td>
<td>Middle</td>
<td>0.24 0.36 0.12</td>
</tr>
<tr>
<td>Top</td>
<td>Top</td>
<td>0.26 0.37 0.13</td>
</tr>
</tbody>
</table>

Partial Reform

Distortion in first age range

Table 15: Intratemporal distortions in Case 4 model simulation

<table>
<thead>
<tr>
<th>Wage quantile in 30-44 age range</th>
<th>Wage quantile in 45-59 age range</th>
<th>Expected value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>Bottom</td>
<td>0.30 0.34 0.08</td>
</tr>
<tr>
<td>Middle</td>
<td>Middle</td>
<td>0.31 0.35 0.09</td>
</tr>
<tr>
<td>Top</td>
<td>Top</td>
<td>0.23 0.29 0.04</td>
</tr>
</tbody>
</table>

Full Optimum

Distortion in first age range
As in all previous models, high-skilled young workers are inefficiently discouraged from working by (here, slightly) higher intratemporal distortions under an age-independent tax system. In this scenario, the inefficiently high distortions due to a lack of age dependence in the Static Mirrlees policy are even more prominent for the high-skilled later in life. Also as before, the use of marginal distortions is lower in the Partial Reform policy than in the Static Mirrlees policy in general, with the expected distortion falling as the policies become more sophisticated, for example from 0.239 to 0.221 to 0.176 across the three policies for the first age group. The smaller magnitude of these results in Case 4 than in the other models is likely due, at least in part, to the compression of the data into two age groups. This compression limits the extent of cross-age incentives problems that cause the Static Mirrlees planner to use more distortionary taxation than the Partial Reform planner.

Age dependence continues to yield a large welfare gain and capture a substantial share of the welfare gain from reform to the Full Optimum policy. Table 16 shows social welfare and lifetime utility for individuals with three representative wage paths under the three policies.

<table>
<thead>
<tr>
<th>Wage paths</th>
<th>Social Welfare</th>
<th>Lifetime Utility (consumption equivalents**)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(consumption equivalents*)</td>
<td></td>
</tr>
<tr>
<td>Always lowest wage</td>
<td>3.908</td>
<td>2.02 2.07 2.23</td>
</tr>
<tr>
<td>Always middle wage</td>
<td>3.921</td>
<td>2.04 2.08 2.22</td>
</tr>
<tr>
<td>Always highest wage</td>
<td>3.940</td>
<td>2.06 2.09 2.20</td>
</tr>
</tbody>
</table>

* The value for consumption that, if provided freely to all workers at each age, would generate the same social welfare as the actual allocation.
** The value for consumption that, if provided freely to the worker at each age, would generate the same lifetime utility as the worker obtains with the actual allocation.

Partial Reform generates a welfare gain equivalent to 0.91 percent of aggregate consumption in the Static Mirrlees. This captures 42 percent of the welfare gain from reform to the Full Optimum. As in previous models, the low-skilled particularly benefit from reform: both the Partial Reform and Full Optimum policies provide more redistribution than the Static Mirrlees policy.

5 Discussion and Special Topics

The preceding sections show that the partial reform of age dependence yields large absolute and relative welfare gains by systematically altering optimal labor income taxation. Moreover, these effects are robust to fundamental variations in the assumed economic environment. In this section, I discuss additional topics of interest that were not addressed directly in these analyses.

5.1 Endogeneity of wage paths

As stated at the beginning of the baseline model setup, I assume throughout this paper that wage paths are exogenous to individuals. Some recent work has begun to address this issue, such as Grochulski and Piskorski (2006) and Kapicka (2003), but the assumption of exogeneity is standard in the optimal tax literature. This does not mean it is innocuous.

If, as this paper’s analysis recommends, tax schedules were to differ by age, individuals would have an incentive to tailor their career choice and employment relationships to minimize their tax bill. This
could reduce the variation in wage distributions with age that gives age-dependent taxes their power and introduce additional distortions to the economy. For instance, lower average tax rates on young workers would encourage people to take jobs with flatter income profiles and to bargain with their employers to shift the timing of income.\footnote{It is important to be clear that while wages are assumed to be exogenous, this paper’s analysis allows income to respond to taxes because individuals choose their level of labor effort.}

The specific results of this paper therefore require that a substantial portion of the variation of wages with age is inelastic to taxes. A few considerations suggest that this requirement’s effects on the paper’s results may be limited.

First, this paper’s focus on age-dependent taxes between the ages of 30 and 60 limits concerns about distorting individuals’ career choices. In their late teens and twenties, individuals have substantial opportunities to shift the timing of higher education and job training to respond to taxes or other incentives.\footnote{Age dependence during this younger age range would be a more treacherous reform to design, though properly-designed age-dependent taxes during this range would potentially add significantly to the welfare gains calculated in this paper.} By age thirty, however, nearly all have completed their education and begun careers, so any distortions to career choice would apply to only those individuals who were substantially forward-looking and for whom the distortion itself had relatively small costs.

Second, the temporary nature of most employer-employee relationships provides a natural barrier to shifting income across ages in response to age dependence, because shifting income is risky without long-term contracts that tie employees to employers. For instance, the taxes recommended by this paper’s analysis would imply shifting income forward, so that some of a worker’s earnings in her forties would be pre-paid to her in her thirties. Of course, an employer will be hesitant to do this unless the worker can commit to remaining at the firm through her forties. Such commitments are rare in modern labor markets.

More generally, different theories of wage determination suggest different sensitivities of income’s timing to taxation.\footnote{For instance, if learning about the quality of matches or on-the-job training is important, individuals are unlikely to be able to shift income to earlier in their careers, while if the acquisition (but not the timing) of outside training is important, it might be shifted to an age at which taxes on income are higher.} The key for this paper is that wages rise over the lifecycle because the passage of time is, for whatever reason, required for a given worker’s effort to be worth more to their employers.

Finally, one piece of evidence suggests that the sensitivity to taxes of both career choice and the timing of income is limited. Currently, wage paths rise sharply with age, especially at high incomes. These upward-sloping paths exist in the context of progressive taxation that ought to encourage flat wage profiles. If wage paths were highly elastic to tax incentives, we would expect to see smoother wage profiles than we do.

Despite the potential importance of wage path endogeneity, characterizing optimal dynamic taxation (and age-dependent taxation) with endogenous wage paths is beyond the scope of this paper and is an important task for future work. Doing so will require a careful treatment of career choice and the timing of income, as discussed above. It will also require the inclusion of a model of human capital investment through education and work experience, a factor that may increase the welfare gains from age dependence as its added flexibility could be used to encourage human capital accumulation.

\subsection*{5.2 Elasticity of Labor Supply by Age}

The analyses of the preceding sections, other than a brief discussion in Section 1, have ignored one of the most direct reasons for the differentiation of taxation by age or any other personal characteristic: variation in the elasticity of labor supply across subgroups (see, for instance, Alesina and Ichino (2007) on gender).
Standard optimal tax theory implies that less elastic subgroups should face larger tax distortions, all else the same, as revenue can be raised more efficiently from them. Therefore, a potentially important determinant of age-dependent taxes absent from this paper’s results is variation in the elasticity of labor supply across ages.

Unfortunately, empirical evidence on variation in the elasticity of labor supply with age is limited. Kremer (2002) argues that "The limited available evidence suggests that younger workers have more elastic labor supply than prime-age workers," citing Clark and Summers (1981), who show more variation in employment rates with the business cycle for young workers. French (2005) estimates that "labour supply elasticities rise from 0.3 at age 40 to 1.1 at age 60," but estimates for other ages are not given. Lacking more robust evidence, I have made the conservative assumption that the elasticity of labor supply is uniform across age.

If labor supply elasticity varies in the directions suggested by this limited evidence, the recommendations of this paper are strengthened. To illustrate this, I consider a parameterization that includes a simple difference in elasticities by age. Consider the baseline model from Section 1 where there are \( I = 10 \) types of individuals living for \( T = 3 \) periods. Suppose that \( \sigma = 3 \) for the second age group (workers in their forties) while \( \sigma = 2 \) for the workers in their thirties and fifties. Given the isoelastic disutility function (9), the constant-consumption elasticity of labor supply is \( \frac{1}{\sigma-1} \), so these values imply an elasticity of 1 for the youngest and oldest groups and an elasticity of \( \frac{1}{2} \) for the workers in their forties.

The results of this experiment are similar to those of the main analyses. Intratemporal distortions remain too high for the high-earning young and are used more in general by the Static Mirrlees policy than by the policies with age dependence. Average tax rates for workers in their thirties are even lower relative to older workers in this Partial Reform policy than in the model with uniform elasticities. The welfare gain from Partial Reform is unchanged at 1.8 percent of aggregate income, and age dependence captures over two-thirds of the potential gain from the Full Optimum.

### 5.3 Extensive Margin

One reason that we may intuitively think the elasticity of labor supply is higher for the young and old is not captured by the previous discussion. Young and old workers may be elastic along the extensive labor supply margin (the choice whether to work or not) rather than the intensive margin (the choice of how much to work). How would an extensive margin affect this paper’s results?

To add an extensive margin to the analysis, I modify the baseline model of Section 1 to include an eleventh type of individual, type \( i = 0 \), who never works. A worker with type \( i > 0 \) who chooses not to work is operating on the extensive margin. Note that, because the Partial Reform planner cannot make history-dependent allocations, workers can move across the extensive margin in a single period or any combination of periods.

To properly model this extensive margin, I must make not working qualitatively different from working less. To do so, I add a fixed cost of working, \( \phi \). Formally, the incentive constraints in the Partial Reform planner’s problem for individual \( i \) of age \( t \) have two parts: first,

\[
\beta^{t-1} \left( u\left(c^t_i\right) - v\left(\frac{y^t_i}{w^t_i}\right) - \phi \right) \geq \beta^{t-1} u\left(c^0_i\right),
\]  

(20)
for all $i \in \{1, 2, \ldots, I\}$ which prevents $i > 0$ from preferring not to work; and second,

$$\beta^{t-1} \left( u(c^t_i) - v \left( \frac{y^t_i}{w^t_i} \right) - \phi \right) \geq \beta^{t-1} \left( u(c^t_i) - v \left( \frac{y^t_j}{w^t_j} \right) - \phi \right),$$

(21)

for all $i, j \in \{1, 2, \ldots, I\}$ and all $t$, which simplify to the same conditions as (6) from the baseline model because $\phi$ cancels on both sides. I simulate the model with $\phi = \ln(1.712)$, representing a fixed cost equal to approximately 30 percent of the log average wage for type $i = 1$.

The lessons from this paper’s main analyses are unchanged by adding an extensive margin, though the optimal policies do respond to the extensive margin. One response of policy is that, while average tax rates have the same shape as in the baseline model, they are increased throughout the income distribution by the addition of an extensive margin. Intuitively, consumption is being provided to the $i = 0$ individuals who do not work, so average taxes on all other types must increase. A second, more subtle response is consistent with the analysis of Saez (2002). In the simulation of policy with an extensive margin, allocations mimic the U.S. Earned Income Tax Credit, whereby low earners receive a subsidy (i.e., a negative marginal tax rate) to encourage them to work rather than claim the $i = 0$ allocation.

5.4 Pareto-improving age dependence

The main analysis in this paper assumes that social planner’s problem is to maximize a Utilitarian social welfare function. This is a restrictive though standard assumption, and concerns about it have inspired research on Pareto efficient taxation such as Stiglitz (1987) and, more recently, Werning (2007b). In a similar vein, the original partial reform approach of Guesnerie (1977) stressed incremental Pareto improvements to tax policy, not incremental Utilitarian improvements. For those uncomfortable with reforms that sacrifice the welfare of some individuals for greater gains by others, the key question is whether age dependence is a Pareto-improving partial reform: that is, a reform that can raise social welfare without harming any individuals.\textsuperscript{47}

Pareto-improving age dependence would also be more likely to succeed as a policy proposal. In particular, concerns about the impact of moving to an age-dependent system can be mitigated by using some of the surplus value generated by the Pareto improvement to compensate those who would otherwise lose in the transition.\textsuperscript{48}

To test whether age dependence is a Pareto-improving partial reform, I simulate the baseline model with the additional restriction that no individual can be worse off under the age-dependent policy than under the Static Mirrlees policy.\textsuperscript{49} As in the Utilitarian model, marginal distortions on high-income young workers and average taxes on all young workers are lower under the Pareto-improving age dependent tax policy than under the Static Mirrlees. More important, the welfare gain from Partial Reform is nearly as large as in the baseline, equivalent to 1.7% of aggregate consumption in the Static Mirrlees. The Pareto-improvement restriction ensures that the highest earners are left with their utility levels from the Static Mirrlees policy, while reform generates a substantial increase in welfare for lower earners. This result suggests that age dependence is a reform capable of attracting broad-based support.

\textsuperscript{47}Blomquist and Micheletto (2003) illustrate the theoretical potential for age dependence to be Pareto-improving.

\textsuperscript{48}Another option to avoid transition concerns is to make age dependence apply only to generations born after the date of the policy being approved.

\textsuperscript{49}Recall that there is no mortality risk in the model economy. In reality, individuals with shorter lives may be relatively disadvantaged, as taxes are likely to be lower on the old. As the welfare benefits calculated above are based solely on the ages between 30 and 59, however, the affected population is small.
6 Conclusion

In this paper, I studied a partial reform of tax policy: age-dependent labor income taxes. To do so, I used modern dynamic Mirrleesian optimal tax methods to contrast three policy scenarios: a Static Mirrlees policy restricted to age-independent taxes, a Partial Reform policy in which labor income taxes can be age-dependent, and a Full Optimum policy in which only private information constrains the design of taxes. In a baseline model, I showed how classic theoretical results on the intratemporal and intertemporal policy margins apply to age-dependent policy. I examined how age dependence affects these margins in economic environments with stochastic wages and private saving and borrowing, as well.

Then, I used data from the U.S. Panel Study of Income Dynamics to calibrate and simulate the three policy scenarios. This quantitative analysis yielded two specific policy recommendations that were largely robust across settings. First, marginal income taxes ought to be lower for high-earning young workers in an optimal age-dependent policy than in an age-independent policy. These individuals are near the top of their age-specific wage distribution, so the efficiency costs of distorting their labor effort are substantial. In an age-dependent tax system, the benefit from such a distortion (increasing tax revenue from higher earners) is relatively small, whereas the benefit appears much larger in an age-independent system that cannot recognize the position of these individuals within their age’s distribution. This specific example illustrates a more general finding that age dependence avoids using marginal distortions that age-independent policy cannot, raising the efficiency of the tax system. Second, younger workers ought to face a lower average tax schedule than middle-aged workers if private saving and borrowing are restricted, as differential average taxes by age substitute for private borrowing in the presence of rising wage paths. In models with private saving and borrowing, a variety of average tax schedules can implement the optimum, including policies that have lower average taxes on the young.

Finally, the calibrated policy simulations allowed me to quantify the welfare gain from age dependence and understand its components. Age dependence yields a large welfare gain equal to between 0.9 and 2.5 percent of aggregate annual consumption. Moreover it captures a substantial portion of the gain from reform to the optimal dynamic policy, ranging from above 40 percent to approximately 95 percent of the potential gain depending on assumptions about the economic environment. Age dependence provides especially large welfare gains for the low-skilled, but most people obtain higher utility than they would under an age-independent policy. In fact, a simulation with the added constraint that age dependence be Pareto-improving yields nearly as large a social welfare gain as does the standard, Utilitarian-optimal age-dependent policy.

These findings show that age dependence, which requires only a simple change to current tax policies, is nevertheless a potentially powerful reform. Future work on age dependence ought to extend this analysis in a few directions outside the scope of this paper.

First, the quantitative analysis of this paper focuses on individuals between 30 and 60 years of age. Some of the largest gains to age dependence may come from individuals outside this range, as wage distributions for people in their twenties and sixties are substantially different from those in the range studied here. This paper neglected those age ranges to avoid large uncertainties about how to treat distortions to the acquisition of human capital early in life (i.e., education) and to the retirement margin in the presence of Social Security, and the age-dependent taxes derived above would have little effect on individuals’ choices on these margins. If these margins were properly modeled, however, the benefits of age dependence may be substantially increased.

Second, as mentioned in Section 5, this paper has assumed that the elasticity of labor supply is constant.
across age groups. While this assumption is almost certainly false, solid evidence on variation in labor supply elasticity with age is surprisingly rare. Substantial variation in this elasticity is likely to raise the value of age dependence, so identifying it should be a high priority for future work.

Finally, an important next step toward taking advantage of this policy opportunity is to use the results of this paper to design and study specific changes to existing taxes. The findings of this paper suggest that such an exercise would identify relatively simple ways to increase the efficiency and equity of current tax policy, yielding substantial welfare gains.

References


