Are Consumption Taxes Really Better Than Labour Income Taxes? Theoretical and Quantitative Implications of the Choice of Tax Base

JOB MARKET PAPER

Kris Shaw*

Department of Economics, University of Pennsylvania

November 11, 2020

Abstract

In many standard economic models, taxes on labour earnings and taxes on consumption are outcome-equivalent. However, this is not the case when taxes are non-linear and households differ with respect to wages and earnings, which is the case considered in this paper. I study the differences between the two tax regimes using a tractable two-period framework and show that the theoretical advantages of consumption taxation are twofold. First, it eliminates an intertemporal distortion on labour supply. Second, consumption is more strongly correlated with lifetime resources, which matters for the distributional impact of the tax system. To assess the quantitative implications of the choice of tax base, I construct a standard overlapping generations model with incomplete markets. After calibrating the model to the U.S. economy, I replace a progressive labour income tax with a progressive consumption tax, taking into account post-reform transition dynamics. This reform produces non-trivial gains in output, consumption and welfare. Most of the benefits stem from improvements in labour efficiency that follow from the mitigation of distortions on work decisions.

Keywords: Optimal taxation, consumption taxation.

JEL Classifications: E62, H21, H24.

*Address: 133 South 36th Street, Philadelphia, PA, 19104, USA, telephone: 1-613-798-6514, e-mail: *shawkr@sas.upenn.edu*.

Contents

1 Introduction

Tax system design has two core problems. One concerns the rate of taxation: should rates vary between individuals or across time? And if so, how? These are critical questions and many fruitful efforts have been made to answer them (Saez, 2001; Kindermann and Krueger, 2017). An even more fundamental question concerns the choice of tax base. *What* should be taxed? After all, fiscal instruments must be selected before they can be calibrated.

This paper focuses on the choice between the taxation of labour income and the taxation of consumption. Specifically, I ask whether there are utilitarian grounds to prefer one base over the other. Several papers in the public finance literature have asked the same question and concluded that consumption taxation is better than labour income taxation (Coleman, 2000; Correia, 2010; Motta and Rossi, 2019). But consumption taxes only deliver efficiency gains in these models because they substitute for missing fiscal instruments. In particular, consumption taxes are used to mimic a levy on initial assets, an inelastic resource that can be taxed without distortion. If tax planners could also use appropriate capital taxes, the differences between the two tax bases would disappear. Because the conclusions of these analyses depend crucially on the exclusion of standard fiscal instruments, one is justified in asking: are consumption taxes really better than labour income taxes?

A more instructive point of departure for this inquiry is Erosa and Gervais (2002). Using a standard life-cycle growth model, they prove that when goverments have access to four fiscal policies—namely debt plus proportional taxes on consumption, labour income and capital income—one of them is redundant. In other words, "it is possible to eliminate either consumption taxes or labour income taxes from a given fiscal policy without affecting the allocation being implemented...This observation applies whether taxes are allowed to be conditioned on age or not."[1]

This equivalency result means that one tax base can dominate the other only in settings that deviate in some way from Erosa and Gervais (2002).[2] Their model features homogenous agents and linear taxes, raising the question of how things might change in environments with heterogeneous agents and non-linear taxes. As Conesa and Krueger (2006) and others have emphasized, progressive taxation plays a potentially beneficial role in such settings by redistributing resources from the rich and lucky to the poor and unlucky. The optimal choice of tax base likely turns on whether one system achieves that objective more efficiently than the other.

To understand the macroeconomic implications of this choice, I study two versions of a

[1]Andres Erosa and Martin Gervais, "Optimal taxation in life-cycle economies", *Journal of Economic Theory* 105, no. 2 (2002): 347.

[2]The aforementioned studies deviate by excluding certain fiscal policies from consideration.

dynamic consumption-savings model, one analytically and the other numerically. The model economies are populated by finitely-lived agents who differ with respect to wages and make endogenous labour supply and savings decisions. In both versions, the tax planner's problem is formulated as a Ramsey-style optimal taxation problem in which the government selects a tax-and-transfer scheme from a given parametric class. The functional form I adopt for the tax code is taken from Benabou (2000) and Heathcote et al. (2017), among others.

My analysis begins with a tractable two-period framework that abstracts from physical capital. Agents draw heterogeneous wage profiles and transfer resources across time using a risk-free bond. I use this model to illustrate the key qualitative differences between the candidate tax structures. The theoretical advantages of consumption taxation are twofold and arise from the fact that wages and earnings fluctuate over time whereas consumption is endogenously smoothed through borrowing and lending. As a result, lifetime resources are more strongly correlated with consumption than with earnings. If the ultimate target of redistribution is lifetime resources, as some writers argue, then consumption becomes an attractive choice of tax base for period-by-period tax systems.[3]

Along similar lines, progressive taxation generates an intertemporal distortion whenever it is linked to a volatile choice variable. To grasp the intuition, consider an agent whose wages change (deterministically) over time. This agent optimally chooses a higher level of work effort in the higher-wage period. This intertemporal effect is dampened, however, if earnings are subject to a progressive labour income tax. Because marginal tax rates increase with earnings, the agent's incentive to tilt hours in the direction of the high-wage period is reduced. Consequently, she flattens her life-cycle labour supply profile and generates lower lifetime earnings. By selecting a relatively smoother base, namely consumption, tax authorities can minimize or even eliminate this type of distortion.

These two differences, which I call the *redistribution channel* and the *efficiency channel*, are not enough to definitely favour consumption taxes over labour income taxes. Changes in the tax base trigger changes in average tax rates to maintain government budget balance. Because these tax rate adjustments have varying impacts on agents of different abilities, a simple conversion of the tax system from an earnings base to a consumption base may not improve welfare. I present a sufficient condition on the underlying wage process that guarantees welfare-superiority of progressive consumption taxation in this environment.

The quantitative model developed in this paper is a richer and more realistic but less

[3]See, e.g., Mirrlees et al. (2011): "The redistributive impact of a tax system is often judged by looking at how much tax individuals pay relative to their income over a relatively short time period–rarely more than a year. But people's incomes tend to change over their lives, which means that this approach can be a poor guide to how progressive the tax system is relative to a person's lifetime income...Ideally, we should judge the distributional impact of the tax system over a lifetime rather than at a point in time." pp. 23-24.

tractable version of the qualitative model. Agents in this economy face idiosyncratic labour market shocks in addition to *ex ante* differences in ability. Yet they still have access to just a single financial instrument—the one-period risk-free bond—which they trade for self-insurance purposes as in Huggett (1993) or Aiyagari (1994). The government finances its expenditures using three sources of revenue: flat taxes on consumption and capital income and a non-linear household tax that is initially assessed on earnings. Additional model ingredients include retirement, mortality risk, accidental bequests and a strict borrowing constraint.

After calibrating the model to the U.S. economy, I perform a series of tax reform experiments. To begin, I quantify the macroeconomic effects of converting the non-linear household tax from a labour income base to a consumption base, holding progressivity constant. By applying the reform only to future generations, I avoid the windfall gains and losses that often complicate welfare assessments along the transition. I find that such a reform leads to non-trivial long-run gains in physical capital (1.9%), output (1.3%) and consumption (1.5%). Most of the reform's benefits stem from improvements in labour efficiency that follow from the mitigation of distortions on work decisions. Using a standard utilitarian welfare criterion, I compute a consumption-equivalent steady-state welfare gain of 0.9%. Because of how I structure the tax experiment, transitional generations experience similar welfare gains to long-run generations. Pre-reform cohorts are subject only to general equilibrium price effects and are largely unaffected. This justifies the use of steady-state comparisons to evaluate the impact of the tax reform.

Because baseline progressivity is almost certaintly sub-optimal for both tax bases, I then perform a best-on-best comparison by numerically characterizing the welfare-maximizing tax code under both regimes, with all tax experiments proceeding along the same lines as the simple reform. Although the utilitarian gap narrows somewhat, the optimal consumption tax still holds a long-run welfare advantage of 0.7% when comparing optima. The main quantitative result, and chief contribution of this paper, still stands: adopting a progressive consumption tax generates moderate welfare gains relative to a tax on earnings.

The paper is organized as follows. The remainder of this section reviews the policy background and the relevant literature. The next part of the paper presents and studies a deterministic and tractable two-period model. Section 2 introduces the basic environment; Section 3 describes the tax design problem and analyzes many of its qualitative properties; and Section 4 discusses some extensions. The paper then shifts to the quantitative analysis. Sections 5 and 6 extend and calibrate the model. Sections 7 and 8 present results from the tax reform experiements. Section 9 discusses sensitivity analyses and Section 10 concludes.

1.1 A Brief Legislative History of Consumption Taxation in the United States

The history of the modern federal income tax begins with the ratification of the Sixteenth Amendment to the United States Constitution in 1913. While income taxes had been collected sporadically throughout the previous century (most notably during the Civil War and its aftermath), the primary sources of federal revenue were tariffs and excise taxes. These indirect levies were usually sufficient to finance the limited activities of the federal government, and in any case there were constitutional obstacles to the adoption of an income tax at the national level.[4] The Sixteenth Amendment cleared these obstacles, paving the way for the government to establish a federal income tax, a power it exercised later that same year with the passage of the Revenue Act of 1913. These acts greatly enhanced the U.S. Treasury's capacity to raise revenue, but also introduced a new hobby-horse for policy-makers and scholars to ride: fundamental tax reform.

A remarkably consistent theme of the tax reform conversation over the past century has been the call to replace income taxes with consumption taxes. Advocates of such reforms have several options to choose from. The simplest version is the retail sales tax, which is levied only on the sale of final goods and services and is remitted to the government by sellers. While a retail sales tax has never been adopted nationally, legislators at the state and local levels are more enthusiastic about its merits, so much so that sales taxes are in force almost everywhere in the country. Including a population-weighted average of local taxes, the combined statutory rate ranges from 0% in four hold-out states (Delaware, Montana, New Hampshire and Oregon), to 9.5% in a trio of southern states (Arkansas, Louisiana and Tennessee).

An alternative form of consumption tax is the value-added tax, or VAT, which is levied on sales of retail and wholesale goods alike. Traders are allowed to deduct the tax charged on their inputs, so the tax effectively applies only to the value added at each stage of production. Because the VAT applies to all sales, there is no need to make legal distinctions between intermediate goods and final goods. There is also a built-in enforcement mechanism as both parties to a transaction are obligated and incentivized to report it to the tax authority. For these reasons, the VAT has proven quite popular internationally. First implemented by France in 1954, a VAT of some kind has been adopted by all OECD countries except the U.S.

[4] After dozens of misfires, Congress successfully passed a federal income tax bill in 1894, but the Supreme Court of the United States struck it down less than a year later in *Pollock v. Farmers' Loan & Trust Company*, 157 U.S. 429 (1895). In the Court's opinion, the income tax violated the constitutional requirement that all direct taxes be apportioned among the states according to population.

A third version is the cash-flow expenditure tax, which, in contrast to the sales tax and the VAT, is a form of direct taxation. The logic of the cash-flow tax is easily grasped once one observes that only two pieces of information are needed to calculate a person's consumption expenditures: (1) the sum of earnings, incomes, transfers and other receipts; and (2) the sum of net contributions to savings and investment. After subtracting the second sum from the first, what remains must equal consumption. The cash-flow tax operates, therefore, very much like the existing income tax, except that taxpayers are allowed a deduction for net savings. Of course, the existing income tax system already has elements of a cash-flow tax as this is the type of treatment granted to pension plans, individual retirement accounts (IRAs) and 401(k) plans.For this reason the status quo is often called *hybrid* system. The conceit of the ideal cash-flow tax is to do away with these special tax-advantaged programs and their limits, restrictions and penalties, and instead allow all savings to be treated in this manner.[5]

An important difference between these various implementations is readily apparent. Whereas the indirect forms (sales tax and VAT) are necessarily proportional (notwithstanding any exempt or zero-rated items), the cash-flow form can be applied at *graduated* rates. Thus, if a policy-maker desires flexibility in setting the progressivity of the tax system, she is likely to favour the cash-flow expenditure tax. And since this paper concerns settings where taxes are allowed to be non-linear, it is the cash-flow form that I have in mind throughout. Unless directed otherwise, the reader should consider 'progressive consumption tax' synonymous with 'cash-flow tax' in what follows.[6]

Several attempts have been made to revise the federal income tax code along the lines of a cash-flow tax. The earliest occurred in 1921 when Republican Congressman Ogden Mills proposed a progressive "spendings tax" as a partial replacement for the existing income tax. Treasury Secretary Henry Morgenthau advanced a similar plan in 1942 in the wake of the country's entry into the Second World War. Both the 1921 and 1942 proposals died in committee, unable to secure a sufficiently broad coalition of support.

Renewed interest in fundamental tax reform during the 1970s precipitated a thorough investigation by the U.S. Treasury Department into the merits and feasibility of progressive consumption taxation. The case for a cash-flow tax was laid out in *Blueprints for Basic Tax Reform*, with economist David Bradford as lead author.[7] The first version of the document

[5]For a detailed description of the operation of a cash-flow tax, see Kaldor (1955), Andrews (1974) or United States Treasury (1977).

[6] Other names for this type of tax include 'expenditure tax' Kaldor (1955), 'spendings tax' Fisher and Fisher (1942), 'spending tax' McCaffery (2002), 'savings-exempt income tax' Domenici (1994), 'consumption-type personal income tax' Andrews (1974), and 'consumed income tax' Goldberg (2013).

[7]Around the same time, a similarly sweeping review of tax policy in the United Kingdom was undertaken by the Institute of Fiscal Studies under the chairmanship of James Meade. Their report, published as *The*

was published in 1977, with a revised edition appearing seven years later in advance of the Tax Reform Act of 1986. The Reagan administration ultimately rejected personal cash-flow taxation, choosing instead to streamline the existing tax system by drastically reducing marginal rates and by eliminating many deductions, exemptions and loopholes. That said, several consumption-tax elements were passed into law during this era. Most significantly, the introduction and refinement of the IRA and the 401(k) between 1974 and 1986 transformed the tax code into the hybrid system we recognize today.

The 1986 reforms were not resilient[8],however, and fundamental tax reform was back at the top of the political agenda within a decade. In April 1995, a bipartisan trio of Senators co-sponsored a bill proposing a progressive federal consumption tax, which they called the Unlimited Savings Allowance or USA Tax (Domenici et al., 1995). But as in 1921 and 1942, their proposal never reached the floor of the House or Senate for a vote.

As documented by Bank (2003), each of these formal proposals for a personal expenditure tax failed for the same reason: an inability to convince either side of the political spectrum that a cash-flow tax represented a worthy compromise of its political aims. Recall that the essential components of the progressive consumption tax are (1) a graduated rate structure; and (2) the exemption of net savings from the tax base. Opponents on the right applaud the latter but object to the former; they prefer a national sales tax or a flat tax on earnings. Opponents on the left applaud the former but object to the latter; they take exception to wealthy but frugal households avoiding their 'fair share' of the tax burden.[9] With both sides unwilling to sacrifice its ideological commitments, the cash-flow tax has not yet gained enough political traction despite bipartisan support among moderates.

1.2 A Brief Intellectual History of Consumption Taxation

The case for consumption taxation has a rich intellectual history, dating back to at least Hobbes and comprising contributions from many notable economists and policymakers. The literature is dominated by two standard arguments, one ethical—consumption taxation is *fair*—and the other economic—consumption taxation is *efficient*. In this subsection, I briefly outline the standard arguments, hopefully lending some context to the consistent popularity of the consumption tax among reformers.

Structure and Reform of Direct Taxation, strongly advised a transition toward a progressive consumption tax.

[8]Michael J. Graetz, "Tax Reform Unraveling", *Journal of Economic Perspectives* 21, no. 1 (2007): 69–90.

[9]During the 1921 congressional committee hearings, Representative William Stevenson asked: "I wonder how [Mills] would think a man like the late Russell Sage was bearing his part of governmental expenses when he was drawing his millions and living on $60 a month or thereabouts, and all of that exempt?" A like-minded person today might find an equivalent exemplar in famously frugal Warren Buffett.

I also explain why these standard arguments are not as convincing as their proponents think. The rhetorical failure stems from a general conflation of consumption taxation and zero capital income taxation. It is not uncommon to find that arguments putatively made about taxes on consumption are in fact just arguments about taxes on capital income. But these are two different fiscal instruments, and it is does not follow that tax systems must be based on consumption even if one accepts that the ideal tax rate for capital income is zero. The ubiquity of this false dichotomy means that much of the literature is orthogonal to the particular merits (or demerits) of consumption taxation. A different kind of analysis is needed, which I set out to provide in the rest of this study.

1.2.1 Fairness

In *The Wealth of Nations*, moral philosopher Adam Smith premised four maxims of a sound system of taxation, the first of which requires that taxes be levied according to a person's ability to pay. That is, each taxpayer's contribution ought to be a function of the resources "which they respectively enjoy under the protection of the state."[10] It is this concern for fairness that motivates many classical discussions of taxation. A central question of these discussions, then, is: what measurable quantity best encapsulates a person's ability to pay, income or consumption? For many thinkers, the answer is consumption.

One version of the fairness argument, which anticipates Smith by over a century, posits consumption as the only equitable basis for assessment because consumption measures a person's withdrawal from society's pool of resources while income measures his contribution. From this standpoint, an income tax is unfair precisely because it penalizes work, wealth-creation and thrift. Importantly, this is not an economic argument about incentives and elasticities. It is not wrong to penalize work, wealth-creation and thrift (just) because they are economic goods; it is (also) wrong because they are *moral* goods. In the words of this perspective's seminal advocate:

> The Equality of Imposition consisteth rather in the Equality of that which is consumed, than of the riches of the persons that consume the same. For what reason is there, that he which laboureth much, and sparing the fruits of his labour, consumeth little, should be more charged, then he that liveth idlely, getteth little, and spendeth all he gets; seeing the one hath no more protection

[10] Smith's other three maxims are: (2) "The tax which each individual is bound to pay ought to be certain, and not arbitrary"; (3) "Every tax ought to be levied at the time, or in the manner, in which it is most likely to be convenient for the contributor to pay it"; and (4) "Every tax ought to be so contrived as both to take out and to keep out of the pockets of the people as little as possible over and above what it brings into the public treasury." See *The Wealth of Nations*, Bk. V, Ch. ii, p. 825-826.

> from the Common-wealth than the other? But when the Impositions are layed upon the those things which men consume, every man payeth Equally for what he useth: Nor is the Common-wealth defrauded by the luxurious waste of private men.[11]

The force of the 'common pool' argument is diminished, however, by the inexorable bind of the taxpayer's budget constraint. Taxpayers cycle through the roles of 'spender' and 'saver' throughout their lives, and status as one or the other is as much a product of age, need and random fluctuations in circumstances as it is of underlying characteristics of taste and temperment. Whatever is earned is eventually spent, and he who augments the Common-wealth today returns tomorrow to deplete it. Because the distinction between earner and spender is fuzzy, this particular notion of fairness may bear less on the question at hand than first supposed.

John Stuart Mill endorsed consumption taxation as the fairest system on somewhat different grounds. His concern lay chiefly with the treatment of savings and investment. For Mill, it was fundamentally unfair that a person be taxed twice on the same part of his resources, once when it was earned and invested, and again when it yields a financial return. Because a future sum consisting of principal and accumulated interest is equal in present-value terms to the principal alone, the taxation of capital income introduces an unjustified bias against savers, no different than if a heavier sales tax were arbitrarily imposed on one kind of widget but not on others. Mill was especially mindful of non-wealthy taxpayers who had no means to provide for retirement or for dependents except by saving out of current earnings. As such, "no income tax is really just from which savings are not exempted."[12] The 'double taxation' story is certainly a popular political argument in favour of consumption taxes. But it is important to note that this concept of fairness pertains to the taxation of capital income; it does little to elucidate the choice between consumption and labour income, which is the topic at hand. Questions about the proper role of capital income taxation fall outside the scope of this paper.

Both Hobbes and Mill endorse consumption as the best measure of a person's ability to pay on moral grounds. A third version of the fairness argument lets revealed preference settle the matter, thereby avoiding abstract moral reasoning altogether. In his seminal proposal for an expenditure tax in the United Kingdom, Kaldor (1955) writes:

> Accruals from the various sources cannot be reduced to a common unit of

[11] Thomas Hobbes, *Leviathan*, Pt. I, Ch. 30, para. 181. Many others have endorsed this logic, including classical economists (see Kaldor (1955) for a summary), treasury secretaries (Summers, 1984) and legal scholars (Goldberg, 2013).

[12] John Stuart Mill, *Principles of Political Economy*, Book V, Ch. 11, section 4.

> spending power on any objective criteria. But each individual performs this operation for himself when, in the light of all his present circumstances and future propsects, he decides on the scale of his personal living expenses. Thus a tax based on actual spending rates each individual's spending capacity according to the yardstick which he applies to himself. (p. 47)

According to Kaldor, the normative case for consumption taxation need not involve any moral judgments about net contributions to collective prosperity (as in Hobbes) or the putative unfairness of double taxation (as in Mill). It is enough that consumption taxation discerns each person's taxable capacity from the decisions she freely makes for herself.

1.2.2 Efficiency

The second standard argument is straightforward: by exempting savings, a consumption-based tax system eliminates intertemporal distortions on saving and investment, thereby encouraging capital accumulation and stimulating economic growth (Seidman, 1989; Frank, 2005; Bankman and Weisbach, 2006; Carroll and Viard, 2012). And tax reform advocates may be correct on this point. But this is not so much an argument *for* consumption taxes as it is an argument *against* capital taxes. Indeed, it is perfectly feasible for a tax system to consist of a tax on capital income in addition to a tax on earnings or consumption. Whether the former is a bad idea or not is an important question (Atkeson et al., 1999; Conesa et al., 2009; Fehr and Kindermann, 2015), but not particularly pertinent to the choice between a labour income base and a consumption base.[13]

The standard efficiency argument, then, has little to say about consumption taxation *qua* consumption taxation, particularly in comparison with labour income taxation. Perhaps

[13] The literature's conflation of consumption taxation with zero capital income taxation is ubiquitious. Some examples:

> "According to proponents, the aim of the [consumption tax] is to promote saving and investment." (Seidman, 1997, p. 1)

> "The distinction between a consumption tax and an income tax...is that an income tax, at least to some extent, taxes the return on savings and investment, whereas a consumption tax does not." (Auerbach and Hassett, 2005, p. 5)

> "The primary difference between the income and consumption tax approaches lies in their treatments of capital income." (Zodrow, 2006, p. 3)

> "The justification for consumption taxes rests on their built-in incentives to save and invest." (Hall and Rabushka, 2007, p. 63)

> "What does normative tax analysis suggest about the case for the choice between consumption taxation and income taxation, or equivalently, the case for taxing capital income?" (Boadway, 2010, p. 11)

because of this, the distinction between the two is often dismissed and sometimes ignored altogether. Instead, they are cast as economically equivalent approaches for achieving savings neutrality. Some tax reformers that couch their arguments in consumption-tax terms end up proposing a straight labour income tax instead, arguing that since the two systems differ only in the timing of tax payments the administratively simpler wage tax is superior (e.g. Hall and Rabushka, 2007).

But are all other things really equal? With homogeneous households, complete markets and linear taxes, perhaps (recall Erosa and Gervais, 2002). The question remains open, however, for settings with heterogenous agents, uninsurable risk or non-linear taxation.

Consider the unique contribution made by Krusell et al. (1996). They augment the neoclassical growth model with household heterogeneity and a political process for endogenously determining linear tax rates on consumption and/or income. Because the decisive median voter is a low-wealth type by assumption, equilibrium tax rates are higher whenever the scope for redistribution is broadest, viz. when taxation is consumption-based. As a result, steady-state output is lower in a consumption tax regime compared with a labour income tax regime.

Another area where the distinction between consumption taxation and labour income taxation has been duly recognized is the treatment of supernormal returns on investment. As noted by Mirrlees et al. (2011), an earnings tax leaves excess returns (and losses) untouched while a cash-flow tax effectively makes the government a silent partner in its taxpayers' investments, enjoying a share of both windfall gains and windfall losses. There is some debate in the literature as to whether investment risk alone breaks the equivalence between labour income taxation and consumption taxation, with some arguing yes (Ahsan, 1989; Ahsan and Tsigaris, 1998) and others no (Zodrow, 1995).

Supernormal returns do not occur in the models studied here since there is no uncertainty with respect to rates of return. Political mechanisms play no role either. Instead of investment risk or political interests, it is the presence of non-linear taxation and idiosyncratic labour market risk that generates differences in outcomes between the two systems. I study the choice between non-linear labour taxes and non-linear consumption taxes independently of the decision to tax capital or not, paying special attention to any effects on the intertemporal substitution of labour effort.

1.3 Other Related Literature

This paper belongs to the Ramsey tradition of tax design, which optimizes tax policy over an exogenously specified set of fiscal instruments. In contrast, the Mirrlees approach restricts neither the tax base nor the shape of the rate schedule. Mirrleesian tax design proceeds by formulating an appropriate social planner's problem, characterizing the constrained-efficient allocation, and then reverse-engineering a tax code to implement it. The inevitable distortions are not imposed from without; they arise endogenously from the trade-off between insurance and incentives. The disadvantage of the Mirrlees approach is practical. The tax systems it recommends are usually quite complicated and generally depend on the entire history of labour earnings (Kocherlakota, 2005).

Although the ad hoc restrictions of Ramsey tax design are theoretically unfounded, they are simple enough to embed into richer models and more easily translated into applied policy advice. It has also been shown in some cases that the Mirrlees solution offers only small welfare gains over a simple Ramsey-style tax code (Heathcote and Tsujiyama, 2017). For these reasons, the Ramsey approach still dominates quantitative public finance, with progress made by exploring increasingly sophisticated tax instruments in increasingly sophisticated environments.[14] Some papers are directly inspired by insights drawn from the Mirrleesian literature. A good example is Kitao (2010), who lets the capital tax rate vary with labour income in accordance with the standard Mirrleesian result that capital taxation and labour supply are negatively correlated.

My paper contributes to ongoing research on the optimal degree of tax progressivity. In several recent contributions to this literature the welfare-maximizing tax schedule is found to be steeply-sloped (Kindermann and Krueger, 2017; Imrohoroglu et al., 2018; Brüggemann, 2020). In contrast, Boar and Midrigan (2020) consider a wider range of possible tax codes and find that a flat income tax combined with a large lump-sum transfer is close to optimal. The cash-flow tax I study in this paper represents a middle ground. It mitigates the distortion on labour supply in a similar fashion to the flat tax but still allows for flexibility in setting the overall progressivity of the system. The value of that additional flexibility will determine the ultimate usefulness of the cash-flow tax as a fiscal policy.

My paper is linked to research that posits lifetime income as the ultimate objective of redistributive fiscal policy and therefore focuses on tax structures that mimic, as much as possible, a direct tax on lifetime resources. One way of achieving this goal is the cumulative assessment method famously championed by William Vickrey (1939, 1947, 1969, 1992). Although he was mainly motivated by considerations of horizontal equity[15] it has been shown that tax-smoothing of this sort is optimal in settings where the disutility of work is isoelastic

[14]Two areas of active research include wealth taxes (e.g. Guvenen et al., 2019; Kaymak and Poschke, 2020) and universal transfers (e.g. Luduvice, 2019; Daruich and Fernández, 2020).

[15](Vickrey, 1939, p. 379): "It has long been considered one of the principal defects of the graduated individual income tax that fluctuating incomes are, on the whole, subjected to much heavier tax burdens than incomes of comparable average magnitude which are relatively steady from year to year."

and wages, while heterogeneous, grow at the same rate for all workers (Werning, 2007; Diamond, 2006). Quantitative assessments of a hypothetical lifetime taxation system include Huggett and Parra (2010) for the U.S. and Haan et al. (2019) for Germany. A tax on annual expenditures like the one studied here approximates a tax on lifetime earnings, at least to the extent that households smooth consumption over time through borrowing and saving.[16]

To the best of my knowledge, this paper is the first to quantify the relative merits of *non-linear* consumption taxation in a model of heterogenous agents and incomplete markets. But several papers have quantified the shift from income taxes to *linear* consumption taxes, including Summers (1981), Auerbach et al. (1983), Altig et al. (2001) and Nishiyama and Smetters (2005). These studies offer two important methodological lessons. First, the analyst must properly account for the induced transition, especially the path of government debt. An increased reliance on consumption taxation in a life-cycle setting tends to increase household demand for assets. But it also increases the government supply of debt, which largely crowds out the accumuation of physical capital. If this is not properly accounted for, the results will grossly overestimate the long-run welfare gains of reform. Second, it is vital that the analyst carefully consider the treatment of existing wealth. A newly imposed consumption tax mimics a lump-sum tax on pre-existing wealth while a labour income tax produces the opposite effect. To insulate the numerical results from bias, it is essential that the reform be designed to minimize windfall gains and losses.

A final strand of literature that merits discussion is the sizable body of research on tax-favoured saving plans like IRAs and 401(k)s. Many empirical studies have documented the impact of these programs on household savings behaviour. A central concern is whether contributions to retirement accounts constitute 'new saving' (as found by Poterba et al., 1995) or merely substitution from pre-existing accounts (as found by Attanasio and DeLeire, 2002). The same question has been investigated in quantitative frameworks by İmrohoroğlu et al. (1998), Love (2006) and Fehr et al. (2008), with similarly conflicting results. Structural models of tax-deferred retirement saving have also been used to identify the extent of hyperbolic discounting, namely by Laibson et al. (1998, 2007). Because tax-favoured retirement accounts shield assets from ongoing taxation on realized returns, but cash-flow taxation does not (at least not in the form I study), my analysis deviates from these other studies by abstracting from changes in saving incentives. Instead, I focus attention on how tax-smoothing impacts intertemporal labour supply and the delivery of insurance through the tax system.

[16]The link between annual consumption and lifetime earnings in the context of tax design has been noted by McCaffery (2005) and Mirrlees et al. (2011), but given little formal attention in the literature.

2 The Qualitative Model

2.1 Basic Environment

Households The economy is populated by agents who live for J periods (and hence by J overlapping generations) and who are endowed with one unit of time at every age. A continuum of new agents is born in each period, each of whom draws an idiosyncratic productivity profile $w = (w_1, w_2, \ldots, w_J)$ from some distribution F. Each generation is identical to the next, meaning that the cross-section of productivities is time-invariant. Because all uncertainty is resolved at the moment of economic birth, each household acts with full knowledge of its future.

Labour is the sole input into a linear production technology. An agent with productivity w who works h hours generates output $y = wh$. Under the assumption of competitive labour markets, an agent's productivity w can be taken as her wage rate (hence the notation).

At each age j, the agent chooses hours h_j and consumption c_j. Households can freely save and borrow subject to a lifetime budget constraint. For simplicity, I set the discount rate and the interest rate equal to zero.

Preferences over streams of consupmtion and leisure are assumed to be time-separable, with a period utility function given by:

$$u(c, h) = \log c - \phi \frac{h^{1+\gamma}}{1+\gamma} \tag{1}$$

The parameter $\phi > 0$ represents disutility of work effort while the parameter $\gamma > 0$ governs the elasticity of labour supply. In particular, the static Frisch elasticity is $\frac{1}{\gamma}$. The household's maximization problem is therefore:

$$\text{(H)} \qquad \begin{aligned} U(w; \hat{T}) = \max_{\{c_j, h_j\}} \quad & \sum_{j=1}^{J} \left(\log c_j - \phi \frac{h_j^{1+\gamma}}{1+\gamma} \right) \\ \text{s.t.} \quad & \sum_{j=1}^{J} (w_j h_j - c_j - \hat{T}(c_j, w_j h_j)) \geq 0 \end{aligned}$$

where $\hat{T}(\cdot)$ denotes the household's tax liabilities, which can depend on consumption and/or labour earnings. Since this is a capital-free economy, there is no capital income and therefore no capital income taxes. Though not essential to the results, the absence of phyiscal capital in the model helps focus attention on consumption taxation as an alternative to *labour* income taxation, independently of how the return to savings is treated.

Government The government implements a tax-and-transfer scheme to accomplish two goals: (1) finance exogenous per capita expenditures g; and (2) redistribute resources between households. The government's motivation is utilitarian: it seeks to maximize average lifetime utility. Although there are overlapping generations, the stationarity of the environment means that each generation is effectively self-contained.

The tax-and-transfer scheme under consideration takes the following form[17]:

$$z' = \lambda z^{1-\tau} \qquad \lambda \geq 0 \quad \tau \leq 1$$

where z and z' are the pre- and post-tax quantities of whatever financial category comprises the tax base, and λ and τ and are fiscal parameters. Taxes are then defined as:

$$T(z) = z - \lambda z^{1-\tau} \tag{2}$$

The parameter τ governs the *progressivity* of the tax. A negative value renders the tax system regressive, while $\tau = 0$ implies a proportional tax with flat rate $1 - \lambda$. I direct my attention, however, toward progressive systems. A tax regime is deemed progressive whenever $\tau > 0$. In this case, marginal tax rates increase with z, a fact that follows from the convexity of the tax function.[18] The tax code becomes confiscatory as $\tau \to 1$: all agents are left with exactly λ regardless of their measured tax base.

The parameter λ scales the tax function and determines the cut-off between those who pay taxes and those who do not. The break-even point is $z_0 = \lambda^{1/\tau}$. All agents with $z < z_0$ receive a positive net transfer.

While the tax function has two parameters, the government can only choose τ freely. It must set λ to satisfy a balanced budget constraint:

$$\mathbb{E}[z] = \mathbb{E}[z'] + g \quad \Rightarrow \quad \lambda = \frac{\mathbb{E}[z]}{\mathbb{E}[z^{1-\tau}]} - \frac{g}{\mathbb{E}[z^{1-\tau}]}$$

The Ramsey problem facing the government is thus to choose τ so that the resulting allocation maximizes a social welfare function derived from the household's value function. Given the government's utilitarian motive, we can write this problem as follows:

$$\text{(R)} \qquad \begin{aligned} \max_{\tau} \quad & \int U(w;\tau,\lambda) dF(w) \\ \text{s.t.} \quad & \lambda = \frac{\mathbb{E}[z]}{\mathbb{E}[z^{1-\tau}]} - \frac{g}{\mathbb{E}[z^{1-\tau}]} \end{aligned}$$

I consider two possibilities for the tax base. First, labour income, so that z and z' are

[17] This functional form is well-established in the public finance literature. Its use dates back to at least Feldstein (1969). For recent examples of its use in models with heterogeneous agents, see Guner et al. (2016), Heathcote et al. (2017) and Wu (2017).

[18] The tax function's first two derivatives are $T'(z) = 1 - \lambda(1-\tau)z^{-\tau}$ and $T''(z) = \lambda\tau(1-\tau)z^{-(1+\tau)}$. Notice that the second derivative is striclty positive for $\tau \in (0, 1)$.

properly thought of as pre-tax earnings and post-tax earnings. Second, consumption, so that z and z' can be thought of as expenditures and consumption.

Before moving on to the analysis, I will describe in fuller detail two special cases of the model, namely the one- and two-period versions, since this is all we need to address the qualitative differences between the two candidate tax bases.

2.2 One Period Version

Consider first the case where $J = 1$. There is no reason to distinguish between consumption taxation and labour income taxation in this case since it is trivially true that expenditures and earnings are equal. But it will prove useful in what follows to develop some notation. The static version of the household's problem is:

$$\begin{aligned} U^1(w; \tau, \lambda) = \max_{c,h} \quad & \log c - \phi \frac{h^{1+\gamma}}{1+\gamma} \\ \text{s.t.} \quad & c = \lambda (wh)^{(1-\tau)} \end{aligned}$$

Substituting for consumption, we can rewrite this as:

$$U^1(w; \tau, \lambda) = \max_{h} \quad \log \lambda + (1-\tau)\log w + (1-\tau)\log h - \phi \frac{h^{1+\gamma}}{1+\gamma}$$

Let $\bar{h}$ denote the solution to this maximization problem. After taking first-order conditions, we easily obtain an expression for the the optimal static labour supply:

$$\bar{h} = [(1-\tau)\phi^{-1}]^{\frac{1}{1+\gamma}}$$

Notice that hours are independent of wages. This should not surprise as it is well known that the income and substitution effects arising from variation in wages exactly offset when utility is logarithmic. By substituting the solution back into the objective function we obtain the associated value function:

$$U^1(w; \tau, \lambda) = \log \lambda + (1-\tau)\log w + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log \phi - 1] \tag{3}$$

It follows that the Ramsey problem in the static environment can be written as:

$$\text{(R1)} \qquad \begin{aligned} \max_{\tau,\lambda} \quad & \log \lambda + (1-\tau)\mathbb{E}[\log w] + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log \phi - 1] \\ \text{s.t.} \quad & \mathbb{E}[w]\bar{h} - \lambda \mathbb{E}[w^{1-\tau}]\bar{h}^{1-\tau} - g = 0 \end{aligned}$$

2.3 Two Period Version

Now consider the case where $J = 2$. Earnings and expenditures are generally not equal period-by-period in this case, which, as we will see, has important implications for the government's choice of tax base.

To begin with, suppose that households are subject to both consumption and labour taxes, each belonging to the parametric class (2). Let λ and τ denote the consumption tax parameters and $\hat{\lambda}$ and $\hat{\tau}$ denote the labour tax parameters. The household's problem is then written as:

$$\begin{aligned} U^2(w;\tau,\lambda,\hat{\tau},\hat{\lambda}) = \max_{x_1,x_2,h_1,h_2} \quad & \log c_1 + \log c_2 - \phi\frac{h_1^{1+\gamma}}{1+\gamma} - \phi\frac{h_2^{1+\gamma}}{1+\gamma} \\ \text{s.t.} \quad & x_1 + x_2 = \hat{\lambda}(w_1h_1)^{1-\hat{\tau}} + \hat{\lambda}(w_2h_2)^{1-\tau} \\ & c_j = \lambda x_j^{1-\tau} \quad j = 1,2 \end{aligned}$$

Here, x_j denotes *expenditure* in period j, which is equal to consumption grossed up to include applicable taxes. Since there is a one-to-one relationship between expenditure and consumption, the problem can be formulated using either x or c as a choice variable.

Letting μ denote the Lagrange multiplier on the budget constraint, the first-order conditions are given by:

$$\begin{aligned} x_j: \quad & 0 = (1-\tau)x_j^{-1} - \mu \\ h_j: \quad & 0 = -\phi h_j^{\gamma} + \mu\hat{\lambda}(1-\hat{\tau})w_j^{1-\hat{\tau}}h_j^{-\hat{\tau}} \end{aligned}$$

It follows immediately that the household's optimal consumption path is (unsurprisingly) constant: $c_1 = c_2 = c$. The intertemporal ratio of hours is:

$$\frac{h_1}{h_2} = \left(\frac{w_1}{w_2}\right)^{\left(\frac{1-\hat{\tau}}{\gamma+\hat{\tau}}\right)} \tag{4}$$

Observe that the allocation of labour effort across time depends on the wage ratio and the elasticity parameter γ (naturally), but also the labour tax parameter $\hat{\tau}$. The consumption tax parameter τ, by contrast, is absent. This marks the first important difference between the two tax regimes.

Some further algebra yields expressions for optimal hours;

$$h_j^* = \left[2(1-\hat{\tau})(1-\tau)\phi^{-1}\left(1 + \left(\frac{w_{-j}}{w_j}\right)^{(1-\hat{\tau})\left(\frac{1+\gamma}{\gamma+\hat{\tau}}\right)}\right)^{-1}\right]^{\frac{1}{1+\gamma}} \qquad j = 1,2 \tag{5}$$

for the household's value function under a pure labour income tax;

$$U^L(w;\hat{\lambda},\hat{\tau}) = 2\log\hat{\lambda} + 2\log\left(\frac{(w_1h_1^*)^{1-\hat{\tau}} + (w_2h_2^*)^{1-\hat{\tau}}}{2}\right) - \phi\frac{h_1^{*(1+\gamma)}}{1+\gamma} - \phi\frac{h_2^{*(1+\gamma)}}{1+\gamma} \tag{6}$$

and for the household's value function under a pure consumption tax:

$$U^C(w;\lambda,\tau) = 2\log\lambda + 2\log\left(\left[\frac{w_1h_1^* + w_2h_2^*}{2}\right]^{1-\tau}\right) - \phi\frac{h_1^{*(1+\gamma)}}{1+\gamma} - \phi\frac{h_2^{*(1+\gamma)}}{1+\gamma} \tag{7}$$

Recall that in the static version of the model, the government's set of fiscal instruments was effectively limited to a single policy parameter, namely τ. In the multi-period model, the government is also free to choose the tax base. Thus, we have two competing Ramsey problems, one that implements a labour income tax and another than implements a consumption tax. The next subsection explores the relative merits of these two approaches.

3 Labour Tax v. Consumption Tax

The fundamental tax design problem in models with heterogeneous agents is that certain relevant information (typically the household's underlying skills, as here) is known only to the agents themselves. Since the government is unable to condition the tax code directly on these hidden exogenous characteristics, it must resort to taxes levied on observable endogenous ones. So, for example, instead of taxing *potential* earnings, the government must settle for taxing *actual* earnings.

The question, then, is this: which observable endogenous quantity should we tax, earnings or consumption? Is there any reason for the tax designer to prefer one base over the other? In the life-cycle model of Erosa and Gervais (2002), the answer is no. If an allocation can be implemented using a labour income tax, then it can also be implemented using a consumption tax, a result that obtains whether or not the tax code is conditioned on age.

But we get a different answer here. Indeed, the main theoretical implication of this paper is that there are good reasons to prefer the consumption base. Two features of the model drive this result. First, underlying productivity varies over the life cycle, resulting in household earnings that fluctuate from period to period. Second, taxation is allowed to be *progressive*, so households face time-varying marginal tax rates whenever they are subject to graduated tax rates on labour income. Since marginal tax rates adversely affect work incentives, this means that an agent's labour supply is most (least) severely distorted when she is at her most (least) productive. By flattening the rewards from work across time, a progressive labour income tax generates a double distortion. Not only is the overall level of effort distorted, but also the allocation of effort over the life cycle.

Fundamentally, the problem with the progressive labour tax is that it applies a graduated rate schedule to a fluctuating tax base. To avoid the resulting intertemporal distortion, the tax designer must either abandon progressivity or find an alternative tax base, one that does not fluctuate from period to period. Fortunately, there is such an alternative, one that is smoothed endogenously by the households themselves: consumption.

And so we arrive at the first reason to favour consumption taxes. By breaking the link between when income is earned and when tax is assessed, a consumption-based tax reduces the distorting effects of progressive taxation, leading to more efficient work decisions and higher lifetime output. This effect is not present in Erosa and Gervais (2002) because their model admits ony linear taxes. These insights are formalized in the following two lemmas.[19]

Lemma 1. *Progressive taxation (that is, $\tau > 0$) distorts the* level *of lifetime labour effort whether it is levied on labour income or consumption. In fact, the severity of the distortion is equal. But a progressive tax on earnings also distorts the* allocation *of effort across time. Thus, a progressive labour tax imposes a greater distortion than a progressive consumption tax.*

Proof. Let v denote the lifetime disutility of labour effort. It is given by:

$$\begin{aligned} v &= \phi \frac{h_1^{*(1+\gamma)}}{1+\gamma} + \phi \frac{h_2^{*(1+\gamma)}}{1+\gamma} \\ &= \frac{\phi}{1+\gamma}\left\{2(1-\hat{\tau})(1-\tau)\phi^{-1}\left[\left(1+\left(\frac{w_2}{w_1}\right)^{(1-\hat{\tau})\left(\frac{1+\gamma}{\gamma+\hat{\tau}}\right)}\right)^{-1} + \left(1+\left(\frac{w_1}{w_2}\right)^{(1-\hat{\tau})\left(\frac{1+\gamma}{\gamma+\hat{\tau}}\right)}\right)^{-1}\right]\right\} \\ &= \frac{2(1-\hat{\tau})(1-\tau)}{1+\gamma} \end{aligned}$$

where I use (5) and the fact that $\frac{1}{1+\frac{a}{b}} + \frac{1}{1+\frac{b}{a}} = 1$. Note that $\hat{\tau}$ and τ are interchangeable in this expression. Thus, both tax types have the same impact on lifetime labour effort. However, only the consumption tax leaves the optimal ratio of hours across time undistored, a fact that follows directly from (4). ■

Lemma 2. *Consider household outcomes under two different tax regimes: a pure labour tax and a pure consumption tax. Suppose both regimes are similarly progressive, i.e., set $\tau = \hat{\tau}$. Without loss, let $w_1 > w_2$. Then:*

1. $h_1^C > h_1^L > h_2^L > h_2^C$
2. $h_1^C + h_2^C < h_1^L + h_2^L$
3. $y^c = w_1 h_1^C + w_2 h_2^C > w_1 h_1^L + w_2 h_2^L = y^L$

[19] Note that an analogous argument can be made *against* the consumption base. Specifically, if consumption fluctuates from period to period (say, because there are shocks to marginal utility) then a progressive consumption tax would distort the intertemporal allocation of consumption spending. But this problem does not arise in the current environment because the optimal consumption path in the model is flat (by construction). Consequently, marginal rates do not fluctuate in equilibrium when taxes are levied on expenditure.

That is, the household works fewer lifetime hours under the consumption tax, but allocates more of them to the high-wage period, leading to higher lifetime output. If wages are constant, change all inequalities to equalities.

Proof. Parts 1 and 2 are corollaries of Lemma 1. Since momentary disutility of labour is convex in hours ($\gamma > 0$) and the hours ratio is steeper under the consumption tax, the stated pattern of hours is the only way for lifetime disutility of labour effort to be equal under the two regimes. Part 3 says that despite lifetime hours being lower under the consumption tax, lifetime output is higher, thanks, of course, to the superior allocation of effort. The proof for this is more involved algebraically and therefore relegated to the appendix. ■

The rationale for favouring consumption taxes over labour taxes strengthens when we remember why we desire progressivity in the first place. Recall the two potentially beneficial functions of progressive taxation. First, to help equalize the distribution of resources between different classes of households (*ex ante* redistribution). Second, to help insure against idiosyncratic household risk in the absence of complete markets (*ex post* insurance). Both these functions are best served by adopting consumption as the tax base.[20]

To see why, it is helpful to think of the tax system as a mechanism in which taxpayers send signals to the government, who in turn assigns tax liabilities based on those signals. For the mechanism to implement an effective *ex ante* redistribution program or an effective *ex post* insurance program, it is essential that these signals be informative about the unobserved characteristics of the taxpayers. Consider the case of a household that generates low annual earnings. What kind of household is this? Is it a poor household in a typical year? Or a rich household in an atypical year? If the latter, is the low productivity anticipated or unanticipated? If unanticipated, is the shock transitory or persisent? It is difficult to answer these questions with just a single data point. *Current* earnings are simply not informative enough; what we need is information about *lifetime* earnings. Indeed, one important lesson from the theoretical literature on dynamic optimal taxation is that constrained-efficient tax codes generally depend on an agent's entire history of labour earnings (see Kocherlakota, 2005).

But such schemes are usually inadmissible in Ramsey-style tax problems, where the planner is constrained not just to a parametric class of tax schedules, but also by the implicit requirement that these schedules be functions of current variables only. Of course, it could be argued that this latter restriction is a feature, not a bug. Actual tax codes often do

[20]Because there are no unanticipated productivity shocks in the current framework, only the *ex ante* redistribution motive technically applies here. But the discussion and reasoning is relevant to both concerns, and certainly both apply in the quantitative model of Sections 5 through 7.

depend only on current variables. In this sense at least, the Ramsey approach reflects the tax planner's problem more closely than the Mirrleesian approach.

Faced with this temporal restriction, the tax planner must determine which *current* tax base embodies more information about a household's *lifetime* tax base. With that in mind, let y denote lifetime earnings and observe that:

$$y = \sum y_j = \sum c_j$$

Because consumption decisions follow a simple smoothing rule, a single observation reveals the entire consumption path, and is therefore *perfectly* informative about lifetime resources. Consequently, assigning tax liabilities period-by-period according to consumption is equivalent to assigning tax liabilities at end-of-life according to lifetime resources. The same cannot be said about earnings. In general, we cannot infer y_2 from y_1 or y_1 from y_2, at least not exactly. Current earnings are only *partially* informative about lifetime resources. These insights are formalized in the following sequence of results.[21]

Lemma 3. *The consumption-based dynamic Ramsey problem is isomorphic to the static Ramsey problem.*

Proof. From (7), the value function for the household under a progressive consumption tax is:

$$U^C(w;\lambda,\tau) = 2\log\lambda + 2\log\left(\left[\frac{w_1 h_1^* + w_2 h_2^*}{2}\right]^{1-\tau}\right) - \phi\frac{h_1^{*(1+\gamma)}}{1+\gamma} - \phi\frac{h_2^{*(1+\gamma)}}{1+\gamma}$$

After some algebra, we obtain:

$$\begin{aligned} U^C(w;\lambda,\tau) &= 2\left\{\log\lambda + (1-\tau)\log(w^C) + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\} \\ &= 2\cdot U^1(w^C;\tau,\lambda) \end{aligned}$$

where w^C is the agent's "pseudo-static" wage:

$$w^C = \left[w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1}{\gamma}}\right]\left[1 + \left(\frac{w_2}{w_1}\right)^{\frac{1+\gamma}{\gamma}}\right]^{\frac{-1}{1+\gamma}} 2^{\left(\frac{-\gamma}{1+\gamma}\right)}$$

The "pseudo-static" wage is the constant wage that is welfare-equivalent to the agent's actual wage profile. In other words, it is her hours-adjusted average wage. Notice that when $w_1 = w_2 = \tilde{w}$, the pseudo-static wage is $w^C = (\tilde{w}+\tilde{w})(1+1)^{\left(\frac{-1}{1+\gamma}\right)}2^{\left(\frac{-\gamma}{1+\gamma}\right)} = \tilde{w}$. In contrast, when $w_1 \neq w_2$, the pseudo-static wage is $w^C > \frac{w_1+w_2}{2}$.

[21] If unanticipated productivity shocks are introduced, consumption is no longer perfectly informative about lifetime resources, merely *more* informative. The basic intuition remains.

The government's Ramsey problem is:

$$
\text{(RC)} \qquad \begin{aligned} \max_{\tau,\lambda} \quad & 2\left\{\log\lambda + (1-\tau)\mathbb{E}[\log w^C] + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\} \\ \text{s.j.} \quad & 2\left\{\mathbb{E}[w^C]\bar{h} - \lambda\mathbb{E}[(w^C)^{1-\tau}]\bar{h}^{1-\tau} - g\right\} = 0 \end{aligned}
$$

This problem is identical to R1 except that w is replaced by w^C. ∎

Lemma 4. *The earnings-based dynamic Ramsey problem is* not *isomorphic to the static Ramsey problem.*

Proof. From (6), the value function for the household under a progressive labour tax is:

$$
U^L(w;\lambda,\tau) = 2\log\lambda + 2\log\left(\frac{(w_1h_1^*)^{1-\tau} + (w_2h_2^*)^{1-\tau}}{2}\right) - \phi\frac{h_1^{*(1+\gamma)}}{1+\gamma} - \phi\frac{h_2^{*(1+\gamma)}}{1+\gamma}
$$

After some algebra, we obtain:

$$
U^L(w;\lambda,\tau) = 2\left\{\log\lambda + (1-\tau)\log\left(\Omega(w,\tau)w^L\right) + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\}
$$

where w^L is the agent's "pseudo-static" wage:

$$
w^L = \left[w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\gamma+\tau}}\right]\left[1 + \left(\frac{w_2}{w_1}\right)^{\frac{(1-\tau)(1+\gamma)}{(\tau+\gamma)}}\right]^{\frac{-1}{1+\gamma}} 2^{\left(\frac{-\gamma}{1+\gamma}\right)}
$$

and Ω is defined as follows:

$$
\Omega(w,\tau) = \left[\left(\frac{1}{2^\tau}\right)\left(w_1^{1-\tau} + \left[w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\gamma+\tau}}\right]^{1-\tau}\right)\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\gamma+\tau}}\right)^{-(1-\tau)}\right]^{\frac{1}{1-\tau}}
$$

The function Ω has the following form:

$$
\Omega^p = 2^{p-1}\left(a^p + b^p\right)(a+b)^{-p} = \left(\frac{a^p + b^p}{2}\right)\left(\frac{a+b}{2}\right)^{-p} \leq \left(\frac{a+b}{2}\right)^p\left(\frac{a+b}{2}\right)^{-p} = 1
$$

where the inequality follows from Jensen's inequality. Thus, Ω is less than unity for all wage paths, holding strictly whenever wages are not constant.

The government's Ramsey problem is

$$
\text{(RL)} \qquad \begin{aligned} \max_{\tau,\lambda} \quad & 2\left\{\log\lambda + (1-\tau)\mathbb{E}\log\left(\Omega w^L\right) + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\} \\ \text{s.j.} \quad & 2\left\{\mathbb{E}[\boxed{}\, w^L]\bar{h} - \lambda\mathbb{E}[(\Omega w^L)^{1-\tau}]\bar{h}^{1-\tau} - g\right\} = 0 \end{aligned}
$$

This problem is not identical to the static analogue. It looks very much like R1 but with w replaced by Ωw^L, except that the Ω is mising in one spot (the red rectangle). ∎

Proposition 1. *A period-by-period tax can replicate a progressive tax on lifetime earnings if it is based on current consumption but not if it is based on current earnings.*

Proof. This follows directly from Lemmas 3 and 4. ■

The analysis so far has yielded two reasons for preferring a consumption base over a labour income base. First, consumption taxes do not distort intertemporal work decisions. Households work fewer lifetime hours, but allocate their efforts more efficiently so that lifetime output is higher. Second, a consumption tax allows the government to assign tax liabilities according to lifetime resources. As a result, the government's redistributive aims can be pursued in a more targeted fashion. On both fronts—incentives and redistribution—the case for a progressive consumption tax appears strong.

Strong, but incomplete. While a tax on consumption is indeed equivalent to a tax on lifetime earnings, a tax on lifetime earnings is not equivalent to a tax on the *path* of earnings. There is information in the parts that is lost in the whole.[22] If a labour income tax manages to exploit that information in some way, then that must be balanced against the aforementioned advantages of consumption taxation.

Lemma 5 shows that under a certain restriction on F, the distribution of wage profiles, it can be shown that switching from an earnings base to a consumption base is welfare-improving.

Lemma 5. *Let Assumption 1 be satisfied (see below). Then a consumption tax with progressivity τ is strictly superior to a labour income tax with progressivity τ, for all $\tau > 0$.*

Proof. Consider again the tax planner's problem under a consuption tax (RC) and a labour income tax (RL). Using the government budget constraint to eliminate λ (leaving just one free tax parameter, viz., τ), we can express social welfare as a function of tax progressivity. Respectively for the two tax bases:

$$V^C(\tau) = 2\left\{\log\left(\frac{\mathbb{E}[w_C]\bar{h} - g}{\mathbb{E}[(w_C)^{1-\tau}]\bar{h}^{1-\tau}}\right) + (1-\tau)\mathbb{E}[\log w_C] + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\}$$

$$V^L(\tau) = 2\left\{\log\left(\frac{\mathbb{E}[w_L]\bar{h} - g}{\mathbb{E}[(\Omega w_L)^{1-\tau}]\bar{h}^{1-\tau}}\right) + (1-\tau)\mathbb{E}\log(\Omega w_L) + \left(\frac{1-\tau}{1+\gamma}\right)[\log(1-\tau) - \log\phi - 1]\right\}$$

Let $\Delta(\tau)$ denote the (halved) difference in social welfare between the two tax regimes given a common choice of τ. Eliminating common terms, we obtain:

$$\begin{aligned}\Delta(\tau) = &\left\{\ \log(\mathbb{E}[w_C]\bar{h} - g) - \log(\mathbb{E}[(w_C)^{1-\tau}]) + \mathbb{E}[\log w_C^{1-\tau}]\ \right\}\\ &- \left\{\ \log(\mathbb{E}[w_L]\bar{h} - g) - \log(\mathbb{E}[(\Omega w_L)^{1-\tau}]) + \mathbb{E}[\log(\Omega w_L)^{1-\tau}]\ \right\}\end{aligned}$$

[22] Recall again the lesson from Kocherlakota (2005) and related literature that optimal taxes in a given period usually depend on the household's labour income in that period *and* all previous periods. There are circumstances in which a cumulative lifetime tax is sufficient (Werning, 2007), but these are not general.

Since $\log(\cdot)$ is an increasing concave function, and $E[w_C] > E[w_L]$, we have:

$$
\begin{aligned}
\Delta(\tau) &> \left\{\log \mathbb{E}[w_C] - \log \mathbb{E}[w_C^{1-\tau}] + \mathbb{E}\log(w_C^{1-\tau})\right\} - \left\{\log \mathbb{E}[w_L] - \log \mathbb{E}[(\Omega w_L)^{1-\tau}]) + \mathbb{E}\log(\Omega w_L)^{1-\tau}\right\} \\
&= \{\log \mathbb{E}[w_C] - \log \mathbb{E}[w_L]\} + \left\{\mathbb{E}\log(w_C^{1-\tau}) - \log \mathbb{E}[w_C^{1-\tau}]\right\} - \left\{\mathbb{E}\log(\Omega w_L)^{1-\tau} - \log \mathbb{E}[(\Omega w_L)^{1-\tau}]\right\} \\
&> \left\{\mathbb{E}\log(w_C^{1-\tau}) - \log \mathbb{E}[w_C^{1-\tau}]\right\} - \left\{\mathbb{E}\log(\Omega w_L)^{1-\tau} - \log \mathbb{E}[(\Omega w_L)^{1-\tau}]\right\}
\end{aligned}
$$

Jensen's inequality and the strict concavity of the log-function implies that both bracketed differences are negative. To go further, recall that $w_C \geq w_L \geq \Omega w_L$ for all w, holding with equality if and only if the wage path is constant. Thus, for any wage path, we can write $\hat{w} = w_C^{1-\tau}$ and $(\Omega w_L)^{1-\tau} = \kappa\hat{w}$, where $0 < \kappa \leq 1$. Adopting this simplified notation:

$$
\begin{aligned}
\Delta(\tau) &> \{\mathbb{E}\log\hat{w} - \log \mathbb{E}\hat{w}\} - \{\mathbb{E}\log(\kappa\hat{w}) - \log \mathbb{E}[\kappa\hat{w}]\} \\
&= \{\mathbb{E}\log\hat{w} - \log \mathbb{E}\hat{w}\} - \left\{\mathbb{E}\log\kappa + \mathbb{E}\log\hat{w} - \log\left(\mathbb{E}[\kappa]\mathbb{E}[\hat{w}] + Cov(\kappa,\hat{w})\right)\right\} \pm \log \mathbb{E}\hat{w} \\
&= \log\left(\mathbb{E}[\kappa]\mathbb{E}[\hat{w}] + Cov(\kappa,\hat{w})\right) - \mathbb{E}\log(\kappa) - \log \mathbb{E}[\hat{w}]
\end{aligned}
$$

If $Cov(\kappa, w_C) \geq 0$ then:

$$
\Delta(\tau) \geq \log\left(\mathbb{E}[\kappa]\mathbb{E}[\hat{w}]\right) - \mathbb{E}\log\kappa) - \log \mathbb{E}\hat{w} = \log \mathbb{E}[\kappa] - \mathbb{E}\log(\kappa) > 0
$$

where the final inequality follows from Jensen's inequality. Under these conditions, a consumption tax with parameter τ is superior to a labour income tax with the same parameter. In fact, there must exist some $\delta > 0$ so that the result goes through as long as $Cov(\kappa, w_C) > -\delta$. □

Proposition 2. *Let Assumption 1 be satisfied (see below). Then an optimal progressive consumption tax is welfare-superior to an optimal progressive labour income tax.*

Proof. Let τ_L^* denote the optimal progressivity of a labour income tax. It follows directly from Lemma 5 that social welfare would be higher under a consumption tax with progressivity τ_L^*. *A fortiori*, an optimized consumption tax must be better than an optimized labour income tax. □

The proof for Lemma 5 relies upon the following assumption.

Assumption 1. $Cov(\kappa, w_C) \geq 0$*, where* κ *and* w_C *are defined as above.*

What does it mean for $Cov(\kappa, w_C)$ to be non-negative? Note that w_C reflects undistorted lifetime earnings capacity while $\kappa \in (0, 1]$ reflects the combined penalty imposed on households with uneven wage profiles when earnings are subject to progressive taxation. The covariance is positive when higher pseudo-static wages are associated with *lower* penalties. Thus, Assumption 1 means that low-wage profiles must not be especially flat compared with high-wage profiles.

Too see why this matters, imagine that there are two classes of households. Low-type households earn the same low wage in both periods. High-type households earn high but variable wages. In this world, low-wage profiles are flat wage profiles, implying that $Cov(\kappa, \hat{w})$ is negative. Assumption 1 is not satisfied.

What happens when the government switches the tax base from labour income to consumption, *keeping tax parameters fixed*? The impact on low types is nil. Since their wages are constant over time, they neither borrow nor save and consume what they earn in each period. Consequently, they are unaffected by the change in the statutory tax base. In contrast, high-type households benefit from improved incentives and insurance, leading to higher pre- and post-tax lifetime earnings. But because tax liabilities tend to fall when the tax base is smoothed—as it is for high types in this example—it is possible if not likely that government revenues will fall. The only way to re-balance the budget without changing τ is to lower λ. And this makes the low types strictly worse off, which is especially bad from a social welfare perspective.

The problem here is that the tax penalty imposed on agents with fluctuating productivity changes the *effective* progressivity of the tax. In particular, when uneven wage profiles tend to be high wage profiles, there are two ways to amplify the redistributive function of the tax code: (1) increase τ; and (2) tax earnings instead of consumption. In principle, the government could maintain budget balance by adjusting τ instead of λ, or by adopting a more flexible tax function. But these strategies, however sensible in practice, would make the problem altogether intractable and preclude any possibility of establishing a result like Lemma 5. Assumption 1, then, should be thought of as a *sufficient* condition to obtain a *stronger-than-needed* result. It bears little if at all on the essential intuition underlying the appeal of progressive consumption taxation.

4 What if Consumption Is Not Constant?

It is no accident that optimal consumption paths in the model are constant; this was by design. A critical reader may wonder if any of the results rely upon this abstraction. In this section, I generalize the model to allow for upward- and downward-sloping consumption paths and show how the tax planner can accomodate for age-varying expenditure patterns when designing a progressive consumption tax.

4.1 The Household's Modified Problem

Let β denote the household's discount factor. In the benchmark model I imposed $\beta = 1$, leading to constant consumption streams; I now relax this assumption and admit any $\beta > 0$. The household is subject to consumption taxation and its maximization problem is:

$$(\mathrm{H}\beta) \qquad \begin{aligned} U^{\beta}(w) = \max_{x_1,x_2,h_1,h_2} \quad & \log c_1 + \beta \log c_2 - \phi \frac{h_1^{1+\gamma}}{1+\gamma} - \beta\phi \frac{h_2^{1+\gamma}}{1+\gamma} \\ \text{s.j.} \quad & x_1 + x_2 = w_1 h_1 + w_2 h_2 \\ & c_j = \lambda x_j^{1-\tau} \quad j = 1, 2 \end{aligned}$$

The solution to the household's problem is characterized by the following Euler equations for hours, expenditure and consumption.

$$h_2 = \left[\left(\frac{1}{\beta} \right) \left(\frac{w_2}{w_1} \right) \right]^{\frac{1}{\gamma}} h_1 \qquad x_2 = \beta x_1 \qquad c_2 = \beta^{1-\tau} c_1 \tag{8}$$

Unlike before, expenditure and consumption now generally follow non-constant paths. Notice that the consumption stream $\{c_1, c_2\}$ is distorted by progressive taxation whenever $\beta \neq 1$, that is, whenever non-constant consumption is optimal. From (8), we have:

$$c_2 = \beta^{1-\tau} c_1 \implies \frac{1}{c_1} \neq \frac{\beta}{c_2} \iff u_{c_1} \neq \beta u_{c_2}$$

This wedge in the household's Euler equation is analogous to the one imposed on labour supply decisions by progressive labour income taxation (see Lemma 1). The underlying problem is the same, viz., the application of a progressive rate schedule to a fluctuating base.

There are fiscal tools, however, that will allow the tax planner to correct the intertemporal distortion of consumption. I focus attention on two remedies: (1) age-dependent taxation; and (2) endogenous tax smoothing. I show that both policies restore the planner's ability to replicate a tax on lifetime resources.

Age Dependence

The idea of conditioning taxes on age is not new. Several papers have touted the merits of age-dependent income taxation, with Weinzierl (2011) being a leading example. The key lesson from this literature is that age dependence allows the tax code to accomodate for how the distribution of skills and wages changes over the life cycle. For example, if the famous "no distortion at the top" result applies, then it should apply separately at each age. But this would be impossible to implement if taxpayers of all ages were subject to the same rate schedule.

A similar principle applies to consumption taxation. If consumption increases over the life cycle (as it would if $\beta < 1$), then a highly-endowed young agent and a modestly-endowed old agent might incur the same expenditures. We might wish to treat them differently for redistributive purposes, but would be unable to do so if constrained by an age-independent system.

Fortunately, it is relatively straightforward to introduce age-dependent average tax rates in this setting. Let τ be universal but allow λ to vary with age. In particular, consider:

$$\begin{aligned} \lambda_1 &= \lambda \\ \lambda_2 &= \beta^\tau \lambda \end{aligned} \tag{9}$$

The following proposition describes how adjusting the tax parameters in this way eliminates the intertemporal distortion.

Proposition 3. *When consumption paths are non-constant, an age-dependent period-by-period consumption tax can:*

1. *Eliminate the intertemporal distortion on consumption; and*

2. *Replicate a progressive tax on lifetime earnings.*

Proof. I relegate the proof to the appendix because it is so similar to earlier analyses. □

4.2 Hybrid Taxation

The tax-base question is meaningful in dynamic settings only because of the implicit restriction to period-by-period taxation. If households were taxed once, at death, then the question would be moot—constrained optimal behaviour ensures that lifetime consumption equals lifetime earnings.[23] In the absence of direct lifetime taxation, the consumption base is preferred because it fluctuates less from year to year. By taxing agents according to a relatively smooth base, the tax planner can still design a system that is progressive with respect to lifetime resources.

But when consumption is less than perfectly flat (as when $\beta \neq 1$), tax liabilities vary from year to year and the tax system begins to diverge from its lifetime ideal. One way to fix this problem is to let taxes depend on age, as documented in the previous subsection. Alternatively, the government can let households smooth their tax liabilities directly. This can be accomplished fairly easily by giving households access to both qualifed and non-qualified tax treatments. Under this regime, households decide for themselves how much of their saving to deduct (and, similarly, how much of their borrowing to include). The tax

[23]Notwithstanding bequests given and received, which are absent in this model.

system would be neither earnings-based nor consumption-based, but rather a *hybrid* of the two.

The household's problem under the hybrid system is written as:

$$
\text{(Hh)} \qquad
\begin{aligned}
U^2(w) = \max_{x_1,x_2,h_1,h_2,s} \quad & \log c_1 + \beta \log c_2 - \phi \frac{h_1^{1+\gamma}}{1+\gamma} - \beta\phi \frac{h_2^{1+\gamma}}{1+\gamma} \\
\text{s.j.} \quad & x_1 + x_2 = w_1 h_1 + w_2 h_2 \\
& c_1 = \lambda x_1^{1-\tau} - s \\
& c_2 = \lambda x_2^{1-\tau} + s
\end{aligned}
$$

Here, the choice variable x is re-interpreted as the household's taxable earnings, that is, earnings less net saving into *qualified* accounts. The new choice variable s denotes net saving into *non-qualified* accounts. The key difference between this model and the benchmark model is that the household's tax smoothing is not constrained by its chosen consumption pattern. The intertemporal budget constraint ensures that all earnings are declared taxable at some point, but the household is free to use non-qualified saving and borrowing to shift consumption across time without triggering any consequences for its tax return.

The following proposition describes how hybrid taxation produces the same benefits as age-dependence.

Proposition 4. *When consumption paths are non-constant, a period-by-period hybrid tax can:*

1. *Eliminate the intertemporal distortion on consumption; and*
2. *Replicate a progressive tax on lifetime earnings.*

Proof. I relegate the proof to the appendix because it is so similar to earlier analyses. □

4.3 Age Dependence v. Hybrid Taxation

While both solutions correct for year-to-year changes in household consumption, hybrid taxation holds two advantages over age dependence. The first is practical. The status quo is, broadly speaking, already a hybrid system. For example, American taxpayers have access to both traditional IRAs (consumption base) and Roth IRAs (labour income base). By using a mixture of instruments, a taxpayer can smooth her tax liabilities over time, at least in part. Reforming the tax system along the lines of the hybrid tax model is therefore a matter of refinement, not overhaul. Age dependence, on the other hand, is largely absent from the actual tax code. Introducing age-conditioned rate schedules into the tax system is likely to prove very difficult, both administratively and politically.

The second advantage lies in hybrid taxation's potential to address problems outside the scope of this simple model. Age dependence only works if consumption varies in a predictable age-related fashion, as it does here. It would not work if consumption varied for idiosyncratic reasons. Suppose, for example, that households are subject to preference shocks: the marginal utility of consumption is especially high in some periods and especially low in others, leading to year-to-year variation in consumption. An age-dependent tax system cannot smooth these sorts of fluctuations. It accounts only for trends, not deviations.

Now consider a household's likely behaviour under hybrid taxation. In periods of especially low expenditures (below trend), the household makes unregistered savings. In so doing they forgo the savings deduction and inflate the current tax base. But since their marginal tax rate is relatively low in such periods the increase in tax is muted. Then, in periods of unusually high expenditures, the household can finance its spending needs by drawing down its unregistered savings. This patten of saving and dis-saving allows the household to 'pre-pay' future tax liabilities when it can exploit more favourable tax rates.

Unanticipated shocks are not the only reason why household expenditures vary from year to year. Spending also fluctuates if households make lumpy purchases of large consumer durables, most notably owner-occupied housing. Imagine, for instance, if households had to include downpayments in their tax base in the year of purchase. This would result in large and unfair spikes in taxes owed. But the problem only arises if all financial assets are treated as qualified accounts. If the household could instead save for its downpayment through non-qualified accounts, it could avoid any concurrent tax consequences when making a purchase.

It is true that the dual-treatment option pushes the system away from a pure consumption tax base. But taxing consumption was never the end, only a means. The goal is to build a period-by-period tax system that can redistribute resources with minimal distortion. Any measure that enables households to smooth tax liabilities over time aids in that regard.

5 The Quantitative Model

In this section I develop a standard incomplete markets model where finitely-lived households supply labour elastically and self-insure against idiosyncratic wage and mortality risk as in Huggett (1993) or Aiyagari (1994). Time is discrete and indexed by $t = 0, 1, \ldots, \infty$. In each period, a single final good is produced according to a neoclassical production function and used for private consumption, investment, and government goods and services. All markets are competitive.

5.1 Households

Demographics The economy is populated by agents who live for at most J periods (and hence by J overlapping generations). A continuum of new agents is born in each period and begins working immediately. Education and training occur prior to economic birth and are not modelled. The working life continues until an exogenously-specificed retirement age $j_R \leq J$, after which the household collects social security benefits. The conditional probability of surviving from age $j-1$ to age j is denoted by ψ_j, with $\psi_{J+1} = 0$. The unconditional probability of surviving to age j is denote by Ψ_j and defined by:

$$\Psi_j \equiv \prod_{k=1}^{j} \psi_k$$

The population is assumed to grow at a constant rate n. By a law of large numbers and the stationarity of the demographic structure, $\frac{n\Psi_j}{\Psi_{j+1}}$ tracks the relative size of adjacent cohorts.

Endowments Agents are endowed with one unit of time in every pre-retirement period, a fraction of which is endogenously devoted to labour market activities. Each unit of work time generates $\rho(j, m, n)$ productivity units, where $m \in \mathcal{M}$ denotes a fixed ability type drawn from a distribution F_m and $n \in \mathcal{N}$ denotes an idiosyncratic stochastic component that follows an age- and type-independent Markov chain with transition matrix $\pi(n'|n)$. Thus, productivity varies across households for three reasons: (1) age, which substitutes for experience; (2) pre-market differences, whether instrinsic or acquired; and (3) unanticipated shocks that (potentially) accumulate over the life cycle. Letting w denote the market price per productivity unit, the household's wage rate is given by $w \cdot \rho(j, m, n)$.

Households are born with zero wealth but receive two types of transfers. Retired households collect a social security benefit, denoted b, while working households receive a bequest, denoted q. The bequest comes from the unintended estates of non-terminal-age decedents, which are appropriated by the government and divided equally among the working-age population.

Preferences Households maximize (expected) utility by choosing consumption c_j and, if not retired, hours h_j at every age. Preferences over stochastic streams of consumption and hours are ordered by:

$$\mathbb{E}\left[\sum_{j=1}^{j_R-1}\beta^{j-1}\Psi_j u(c_j,h_j)+\sum_{j=j_R}^{J}\beta^{j-1}\Psi_j\tilde{u}(c_j)\right] \tag{10}$$

where β denotes the common discount factor and u and $\tilde{u}$ denote the period utility functions for working years an retirement years, respectively.

5.2 Government

The government raises revenue to finance exogenous expenditures G, service its accumulated debt B, and fund a social security system that delivers a benefit b to each retired household. It does so by levying taxes and issuing new debt. Its set of fiscal instruments includes linear taxes on both consumption and the return to capital, denoted τ_c and τ_k respectively. It also operates a non-linear tax-and-transfer scheme based on labour earnings, denoted $\hat{T}(\cdot)$. No part of the tax code can be conditioned on taxpayer age.

Finally, as alluded to above, the government collects accidental bequests and redistributes them among the working-age population in a lump-sum fashion. The exclusion of retirees from the spoils of unspent nest eggs is a crude way of accounting for the age distribution of beneficiaries.

5.3 Markets

Output Market A representative firm produces the economy's only good by operating a constant returns-to-scale technology. The aggregate production function is:

$$Q = F(K,N) = AK^{\alpha}N^{1-\alpha} \tag{11}$$

where Q, A, K and N denote the aggregate levels of output, technology, capital and effective labour. Capital earns an output share of $\alpha \in (0,1)$ and depreciates at rate $\delta > 0$.

Factor Markets Spot markets exist for capital and labour, with prices denoted by r and w, respectively.

Asset Markets A representative intermediary trades the economy's sole financial asset: a one-period risk-free bond that pays an interest rate i or, equivalently, a gross rate $R \equiv 1+i$. This intermediary supplies capital to the representative firm and facilitates the intertemporal transfer of resources for both households and government. Notably, there are no assets with which the household can explicity insure against idiosyncratic wage risk or the uncertainty of survival/death. Moreover, the scope of self-insurance is limited by a stringent borrowing constraint applied to all households. The government, on the other hand, is free to borrow and limited only by its ability to pay back debts.

5.4 Equilibrium

Household's Problem I formulate the household's problem recursively. Individual state variables are age j, assets a, ability type m, and current productivity n. Let $z = (j, a, m, n)$ denote the agent's state and Φ_t denote a probability measure describing the distribution of individual states at time t. Given a sequence of prices and policies, the Bellman equation for the working household ($j < j_R$) is:

$$
\text{(H1)} \qquad \begin{aligned} v_t(j,a,m,n) = \max_{h,c,a'} \quad & u(c,h) + \beta\psi_{j+1} \int v_{t+1}(j+1,a',m,\tilde{n})\pi(\tilde{n}|n)d\tilde{n} \\ \text{s.t.} \quad & (1+\tau_{c,t})c + a' = Y_t^d(w_t\rho(j,m,n)h) + R_t a + q_t \\ & c, a' \geq 0, \;\; h \in [0,1] \end{aligned}
$$

where $Y_t^d(y) \equiv y - \hat{T}_t(y)$ denotes after-tax earnings. The Bellman equation for a retired household ($j \geq j_R$) is simpler:

$$
\text{(H2)} \qquad \begin{aligned} \tilde{v}_t(j,a) = \max_{c,a'} \quad & \tilde{u}(c) + \beta\psi_{j+1}v_{t+1}(j+1,a';\Phi_{t+1}) \\ \text{s.t.} \quad & (1+\tau_{c,t})c + a' = R_t a + b_t \\ & c, a' \geq 0 \end{aligned}
$$

Firm's Problem The representative firm hires capital and labour to maximize profits, which are zero in equilibrium by construction. Its optimality conditions are:

$$
\begin{aligned} r &= \alpha A(N/K)^{1-\alpha} - \delta \\ w &= (1-\alpha)A(K/N)^{\alpha} \end{aligned} \tag{12}
$$

Intermediary's Problem The competitive markets assumption implies zero profits for the financial intermediary. Combined with a no-arbitrage condition, this gives:

$$R \equiv 1 + i = 1 + (1 - \tau_k) r \tag{13}$$

Government Budget Government budget deficits (surpluses) are absorbed by increases (decreases) in the stock of public debt.

$$\begin{aligned} G_t + \sum_{j=j_R}^{J} \left(\prod_{k=1}^{j} \psi_k \right) b_t + R_t B_t = \int \hat{T}_t(w_t \rho(j,m,n) h_t(j,a,m,n)) d\Phi_t + \tau_{c,t} \int c_t(j,a,m,n) d\Phi_t \\ + \tau_{k,t} r_t K_t + B_{t+1} \end{aligned} \tag{14}$$

Accidental Bequests Estates must be assigned in order to close the model. I assume that they are collected and redistributed in full to the working-age population.

$$q_t = \begin{cases} \left(\sum_{k=1}^{j_R - 1} \Psi_k \right)^{-1} \int \Psi_{j-1}(1 - \psi_j) R_t a_0 d\Phi_0 & \text{if } t = 0 \\ \left(\sum_{k=1}^{j_R - 1} \Psi_k \right)^{-1} \int \Psi_{j-1}(1 - \psi_j) R_t a_t(j,a,m,n) d\Phi_{t-1} & \text{otherwise} \end{cases} \tag{15}$$

Market Clearing The market-clearing conditions for factor markets are:

$$\begin{aligned} K_t &= \begin{cases} \int a_0 d\Phi_0 & \text{if } t = 0 \\ \int a_t(j,a,m,n) d\Phi_{t-1} & \text{otherwise} \end{cases} \\ N_t &= \int \rho(j,m,n) h_t(j,a,m,n) d\Phi_t \end{aligned} \tag{16}$$

Letting $C_t = \int c_t(j,a,m,n) d\Phi_t$, the resource constraint is:

$$C_t + K_{t+1} + G_t = A K_t^{\alpha} N_t^{1-\alpha} + (1 - \delta) K_t$$

although this can be safely ignored by Walras' law.

Definition 1. Index time by $t = 0, 1, \ldots, \infty$. Fix a sequence of government expenditures $\{G_t\}$ and initial conditions B_0 and Φ_0. A **recursive competitive equilibrium** is a sequence of value functions $\{v_t\}$, allocations $\{h_t, c_t, a_{t+1}, K_t, N_t\}$, prices $\{w_t, r_t, i_t\}$, government policies $\{\hat{T}_t, \tau_{c,t}, \tau_{k,t}, b_t, B_t\}$, transfers $\{q_t\}$, probability measures $\{\Phi_t\}$, and laws of motion $\{H_t\}$ such that for all t:

1. Given prices, policies and transfers, $\{v_t\}$ solves problems H1 and H2, with $\{h_t, c_t, a_{t+1}\}$ being the associated decision rules;

2. Given prices, the firm's allocation $\{K_t, N_t\}$ satisfies (12);

3. Prices $\{r_t, i_t\}$ satisfy the no-arbitrage condition (13);

4. The government budget (14) is balanced;

5. Estates are redistributed in their entirety, as per (15);

6. Markets clear, as per (16); and

7. The distribution over individual states evolves according to $\Phi_{t+1} = H_t \Phi_t$, where the aggregate laws of motion $\{H_t\}$ are consistent with the decision rules $\{h_t, c_t, a_t\}$ and the transtion matrix π.[24]

Definition 2. A **stationary recursive competitive equilibrium** is a recursive competitive equilibrium where the distribution over individual states is stationary. That is, $\Phi_{t+1} = \Phi_t$ for all $t = 1, \ldots, \infty$. A stationary equilibrium is also called a **steady state** of the economy.

6 Calibration

The quantitative analysis begins by making functional form assumptions and choosing values for model parameters. There are two distinct sets of parameters. Externally calibrated parameters are taken directly from other sources or can be estimated independently of the model. Internally calibrated parameters are selected so that model-generated data match a certain set of targets. Although each internally calibrated parameter is associated with a particular target, it is important to keep in mind that they are jointly determined. All parts of the calibration exercise proceed under the assumption of a stationary equilibrium.

Demographics One model period corresponds to one year. Agents enter the economy at age 23 ($j = 1$), retire at age 65 ($j_R = 42$), and die no later than at age 95 ($J + 1 = 73$). The Social Security Administration's Actuarial Life Table for men[25] is used to determine age-dependent survival probabilities $\{\psi_j\}_{j=1}^{J}$, with ψ_{J+1} set to zero. The population growth rate is 1.1%, the long-run averge value for the USA.

Preferences I assume the following functional form for the period utility function:

$$u(c, h) = \frac{c^{1-\sigma}}{1-\sigma} - \varphi \frac{h^{1+\gamma}}{1+\gamma} \tag{17}$$

The parameters σ, φ and γ govern risk aversion, disutility of work, and the intertemporal elasticity of substitution for labour supply. I set $\sigma = 2$ and $\gamma = 2$. Both values are standard

[24] For a formal characterization of this statement, see any number of papers (co-)authored by Dirk Krueger.

[25] http://www.ssa.gov/OACT/STATS/table4c6.html.

in the literature.[26] Disutility of work φ is chosen so that working households devote, on average, one third of their time endowment to labour market activities. The final preference parameter is the discount factor β, which is chosen to generate an equilibrium capital-output ratio of 2.75 or, equivalently, an interest rate of 4.0%. Given additive separability of consumption and hours, the retirement utility function is naturally defined as $\tilde{u}(c) = u(c, 0)$.

Productivity Recall that household productivity depends on age, pre-market differences and unanticipated shocks. With that in mind, consider the following specification for the evolution of the agent's productivity.

$$\begin{aligned} \ln e(j,m,n) &= \beta_{0m} + \beta_{1m}(j-1) + \beta_{2m}(j-1)^2 + \zeta(n) \quad && \beta_m \sim \mathcal{N}(\beta, \Sigma) \\ \zeta(n) &= \eta_j + \varepsilon_j && \varepsilon_j \sim \mathcal{N}(0, \sigma_\varepsilon^2) \\ \eta_j &= \phi\eta_{j-1} + \nu_j && \nu_j \sim \mathcal{N}(0, \sigma_\nu^2) \quad |\phi| < 1 \\ \eta_0 &= 0 \end{aligned} \tag{18}$$

Ability types are associated with a particular coefficient vector $\beta_m = (\beta_{0m}, \beta_{1m}, \beta_{2m})$, with each being an independent draw from the same distribution. Ability, therefore, governs the shape and scale of the household's expected wage profile. The covariance matrix Σ is unrestricted, allowing for correlation between the components of β_m. Let σ_i^2 denote the diagonal elements of Σ and σ_{ij} denote the off-diagonal elements.

The model is silent about why productivity varies from person to person *ex ante*. We might think that the dispersion in intercepts is due to pre-market activities that affect initial human capital (e.g. schooling choice and family background). Similarly, inate differences in the ability to learn and acquire additional human capital might account for variation in the slope and curvature of the wage profile as in Huggett et al. (2011). I abstract from such considerations here and take these *ex ante* differences as exogenous.

The stochastic process linked to state variable n has two components, a persisent shock η_j that follows an AR(1) process and a transitory shock ε_j. Together, these two shocks generate random fluctuations around a deterministic trend. Assume that the innovations are independent of each other and β_m.

I use data from the Panel Survey of Income Dynamics (PSID) to estimate (18) using adjusted male log-wages as the dependent variable. The results are reported in Table 1 and the details of the sample selection and estimation are discussed in Appendix B. The estimates are very precise except for the covariance between the person-specific intercept and the person-specific experience coefficients. The results suggest that, on average, real wages

[26] The choice of γ implies an after-tax Frisch elasticity of 0.5, as recommended by Chetty et al. (2011). More recently, Blundell et al. (2016) report Frisch elasticity estimates of 0.68 for men and 0.96 for women. See Keane (2011) for a survey of the earlier microeconometric literature

grow at about 4% per year early in a person's working life. There is, however, evidence of significant heterogeneity. A person whose β_{1m} falls one standard deviation above (below) the mean will experience initial wage growth of 7% (1.5%) *ceteris paribus*. Over time, these growth rate differentials can generate substantial inequality as initial advantages and disadvantages accumulate over the life-cycle. The effect is partly mitigated by the strong negative correlation between the linear and quadratic coefficients (β_{1m} and β_{2m}).

Table 1: Income Process Estimation Results (log-wages)

Parameter	Estimate	Std Error	t-statistic	P-value	$CI_{0.025}$	$CI_{0.975}$
			Common Trend			
β_0	2.1809	0.1438	15.17	0.00	1.8991	2.4628
β_1	0.0415	0.0007	59.81	0.00	0.0402	0.0429
β_2	-0.0008	0.0000	-52.65	0.00	-0.0009	-0.0008
			Heterogeneous Trend			
σ_0	0.2175	0.0239	9.10	0.00	0.1639	0.2602
σ_1	0.0265	0.0021	12.74	0.00	0.0220	0.0303
σ_2	0.0006	0.0000	13.67	0.00	0.0005	0.0007
$corr_{01}$	0.1475	0.2052	0.72	0.47	-0.1534	0.7687
$corr_{02}$	-0.4588	0.2585	-1.77	0.08	-1.2076	-0.0658
$corr_{12}$	-0.8749	0.0224	-39.05	0.00	-0.9106	-0.8148
			Stochastic Trend			
ϕ	0.7869	0.0224	35.19	0.00	0.7390	0.8301
σ_ν	0.2096	0.0117	17.89	0.00	0.1872	0.2345
σ_ε	0.3884	0.0096	40.53	0.00	0.3681	0.4065

The first-stage regression includes 4286 individuals and 67,009 person-year observations.
The second-stage regression includes 932 empirical variance-covariance moments.

The estimate of ϕ indicates moderate persistence. Roughly half of a persistent shock's effect remains after three years. At the ten year mark, over 90% of the shock has dissapated. This level of persistence is significantly lower than estimates obtained for models that exclude *ex ante* growth rate heterogeneity. Such estimates are typically in the 0.96-1.00 range.[27]

I use a quadrature-based method to approximate both shock processes with discretized Markov chains. I use seven states for the persistent component and three states for the transitory component, meaning that n_j takes on one of twenty-one values in the finite set $\mathcal{N}$. The invariant distributions of these discretized Markov chains are summarized in Table 2.

[27] See, for example, Abowd and Card (1989), Storesletten et al. (2004) or Karahan and Ozkan (2013).

Similarly, I approximate the distribution of wage profiles by selecting a finite number of equiprobable types, each of which is characterized by a vector β_m. Constructing the set of types is complicated by the fact that the coefficients are correlated. There are three possible values for the intercept (think: high, medium, low), three possible values for the slope (conditional on the intercept), and three possible values for the curvature (conditional on both the intercept and the slope). Hence, $\mathcal{M}$ has twenty-seven elements. The details of both discretization procedures are discussed in Appendix C.1.

Table 2: Invariant Distributions of Productivity Shocks

State	Value	Π
η_1	0.3916	0.0025
η_2	0.5534	0.0499
η_3	0.7493	0.2447
η_4	1.0000	0.4058
η_5	1.3345	0.2447
η_6	1.8070	0.0499
η_7	2.5539	0.0025

Persistent shock process.

State	Value	Π
ε_1	0.7712	0.1932
ε_2	1.0000	0.6135
ε_3	1.2967	0.1932

Transitory shock process.

Government Policy The main policy of interest is the non-linear tax on labour income, which takes the following log-linear form by assumption:

$$\hat{T}(y) = y - \lambda y^{1-\tau}$$

I set $\tau = 0.136$ as estimated by Kaplan (2012) and choose λ to balance the government's state-steady budget.[28] The capital income tax rate is set to $\tau_k = 28.3\%$ as in Kindermann and Krueger (2017). The linear consumption tax rate is set to $\tau_c = 4.4\%$, which is equal to the sum of general and selective sales taxes collected by all state and local governments divided by total nominal personal consumption expenditures in 2017.

With respect to the other side of the public ledger, I choose exogenous government spending such that it accounts for 17% of total output. I then take the implied *level* of governent expenditures as fixed in all ensuing tax reform experiments. I choose social security benefits b to equal 35% of average earnings and the stock of outstanding public debt B to yield an equilibrium debt-to-output ratio of 0.97, as in the data.

The government is also responsible for collecting and redistributing accidental bequests. Though not a parameter, q is an equilibrium object that must be solved for within the model.

[28] Wu and Krueger (2019) estimate the same labour income tax function and obtain a similar estimate of $\tau = 0.133$.

Technology Three parameters (A, α, δ) characterize the production technology. I set the capital share to $\alpha = 0.33$ and the depreciation rate to $\delta = 8\%$. The level of technology A is normalized so that the equilibrium wage rate is $w = 1$.

Summary The key non-productivity parameters are summarized in Table 3, with internally calibrated parameters in bold.

Table 3: Benchmark model parameters

Parameter		Value	Target
Preferences			
σ	Coefficient of relative risk aversion	2.0	
γ^{-1}	Frisch elasticity	0.5	
φ	**Disutility of labour**	**63.3**	$\bar{h} = 1/3$
β	**Annual discount factor**	**0.992**	$r = 4\%$
Production			
α	Capital income share	0.33	
δ	Depreciation rate	8%	
A	**Aggregate technology**	**0.937**	$w = 1.0$
Policies			
τ	Labour tax progressivity	0.136	
λ	**Labour tax level**	**0.666**	$G/Q = 0.17$
τ_k	Capital tax rate	28.3%	
τ_c	Sales tax rate	4.4%	
B	**Public debt**	**27.1**	$B/Q = 0.97$
b	**Social security benefit**	**0.198**	$b/\bar{y} = 0.35$
Transfers			
q	**Accidental bequest**	**0.033**	fixed point

Internally calibrated parameters are in bold.

Brief Remark on Computation Calibration requires the repeated solution and simulation of the model. To solve for the optimal decisions rules, I use the endogenous grid method, augmented for endogenous labour supply as in Barillas and Fernández-Villaverde (2007). The chosen functional forms for preferences and taxes mean that no root-finding procedures are needed, except when simulating the decisions of a borrowing-constrained household. See Appendix C for a detailed description of the algorithms employed.

Life-Cycle Profiles in the Baseline Economy Figure 1 displays average hours and earnings profiles for working-age households and average consumption and asset profiles for all households. Assets rise over the working life as household build buffer stocks to hedge against wage shocks and accumulate wealth to finance retirement consumption. The consumption and asset profiles are also affected by a high degree of patience: the subjective discount rate is 0.8% while the after-tax rate of return is substantially higher at 2.9%.

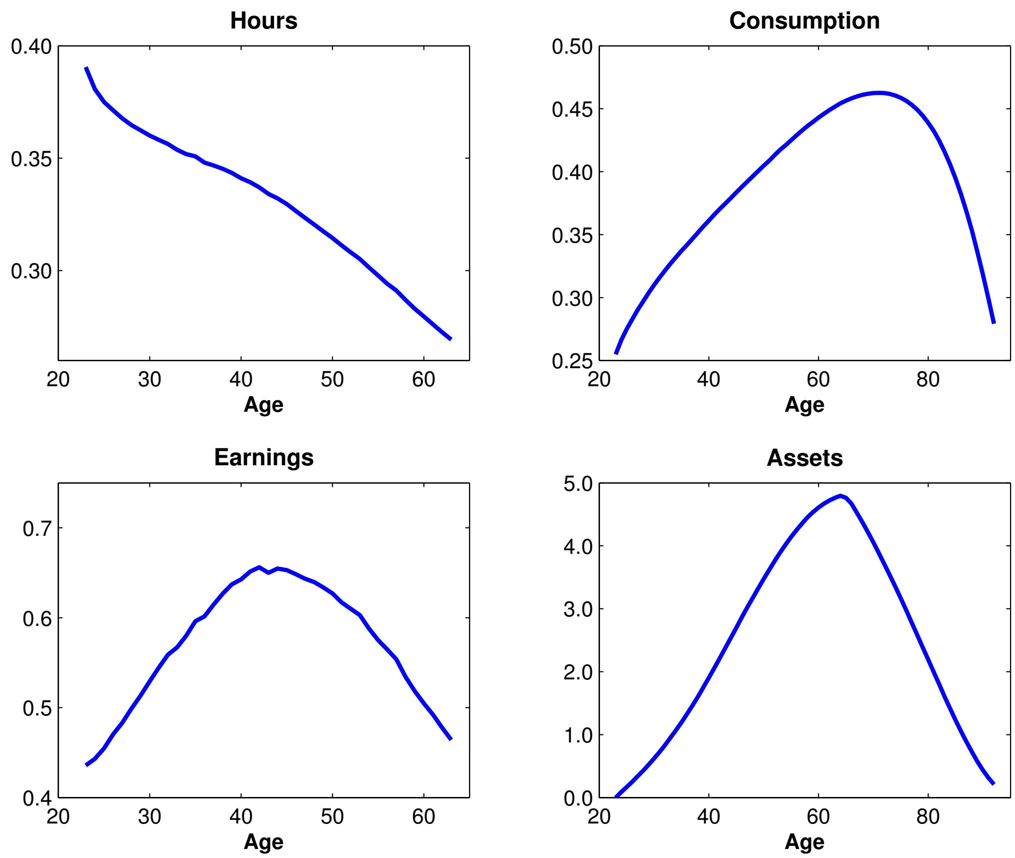

Figure 1: Cross-Sectional Means over the Life Cycle

For reasons explained in Section 3, the degree to which consumption and labour income fluctuate from year to year will be an imporant factor in the comparative assessment of the two candidate tax bases. To that end, I simulate a large number of histories and compute household-specific standard deviations for those two variables. Retirement benefits are counted as labour income for the purposes of this exercise, and all calculations are weighted by survival probabilities and normalized. The densities of these standard deviations are plotted in Figure 2. It is clear from the graph that the model produces considerably less year-to-year variation in consumption relative to earnings.

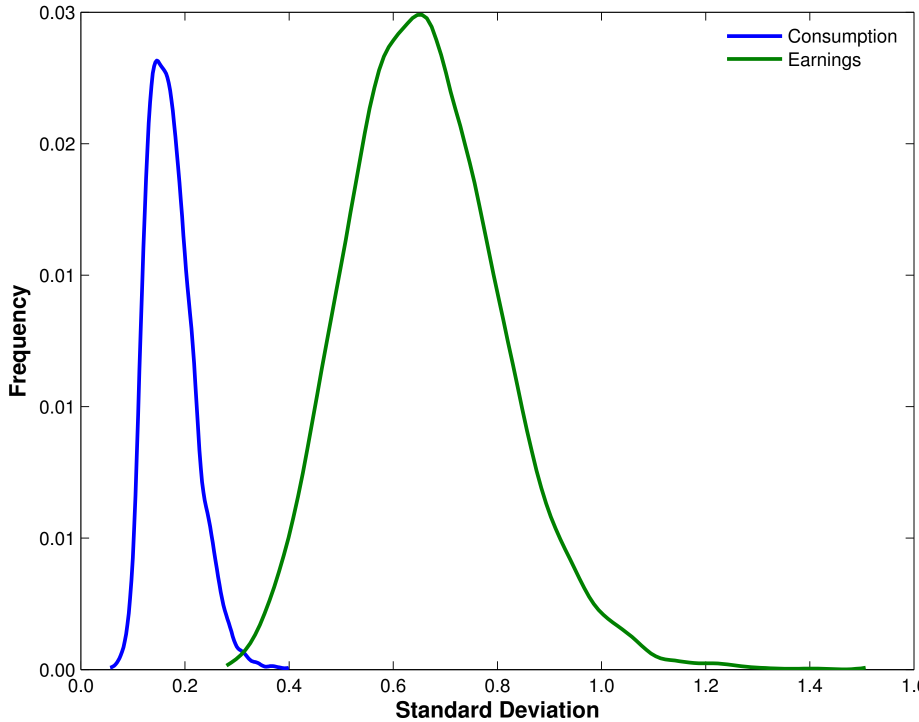

Figure 2: Fluctuations in Consumption and Earnings over the life Cycle

7 A Simple Tax Reform

In this section I use the calibrated model to study a simple tax reform that converts the household tax base from labour income to consumption. I begin by describing the experiment in detail, and then set out the main results. I focus on long-run impacts by comparing stationary equilibria, but also discuss the macroeconomic and welfare consequences along the transition path. The simple reform generates moderate welfare gains by tempering the distortion on household labour supply responses to productivity shocks.

7.1 Description of the Experiment

The economy is in its initial stationary equilibrium at time $t = 0$ when the government announces its plan to convert the household tax-and-transfer system to a consumption base. Progressivity, indexed by the parameter τ, remains fixed though the average tax rate can and will adjust to maintain budget balance. As is standard in the literature, I assume that the policy change is unexpected and the government can credibly commit to making no further changes. Another standard practice I adopt is to fix the *level* of government purchases, allowing the fraction of output devoted to government goods to deviate from its calibrated target.

In practice, this reform is accomplished by allowing a deduction for net savings in the style of IRAs or 401(k) plans. The household's tax base decreases one-for-one with every dollar it contributes to a savings account and increases one-for-one with every dollar it withdraws. Because consumption equals earnings less net saving, this policy successfully implements the intended reform. Unlike IRAs and other real-world tax-deductible savings plans, investments are not sheltered in whole or in part from ongoing capital income taxation. Neither do I impose any limits, restrictions or penalties.

I rule out changes to the capital income tax rate τ_k for two reasons. First, my goal is to isolate and assess a particular policy choice: that between the taxation of earnings and the taxation of consumption. Because capital taxes have important efficiency and distributional impacts (Fehr and Kindermann, 2015), varying τ_k would add unnecessary noise to the results. Second, unanticipated changes to the capital tax structure mimic a lump-sum tax (or transfer) on the existing capital stock. By levying a tax on an inelastic resource, the government can reduce or even eliminate the need for distorting taxes, thereby generating substantial welfare gains from reform. But there are unmodelled political reasons why a fiscal policy of this sort might be unfeasible. One could imagine, for example, that governments are subject to some sort of commitment constraint that prohibits new taxes on old choices. Or, more simply, it could be difficult to garner political support for a policy that punishes living citizens, who can vote, in order to reward yet-to-be-born citizens, who cannot.

A new tax on consumption produces a similar effect. Households that had accumulated assets under the belief that those assets could be liquidated tax-free in the future now face unanticipated tax bills. One way to neutralize this effect is to assign compensatory transfers, as in Nishiyama and Smetters (2005). While this approach provides a useful theoretical measurement of a policy's efficiency effects, it would be difficult to implement in practice, fraught as it would be by information frictions, equilibrium effects, and administrative costs. I adopt an alternative approach, one that is both feasible and easy to imagine in practice. Specifically, I assume that the new tax structure applies only to *newborn* generations, that is, those born at time $t = 1$ or later. Existing cohorts continue to work and save under the rules of the status quo. In this way, they are shielded from the adverse effects of unexpected wealth levies and exposed only to general equilibrium effects on factor prices.

A further detail concerns the balancing of the government budget. One way to do this is to adjust λ in every period along the transition path, thereby keeping the budget balanced period-by-period. This approach, however, produces significant inter-cohort redistribution. A notable feature of consumption is that its time path is delayed relative to the time path of earnings. This is a mechanical consequence of the life-cycle savings motive. Households save when young and productive so that they can consume when old and unproductive. As a result, a consumption tax tends to postpone tax liabilities until later in the life cycle. But this means that the aggregate tax base drops significantly during the early years of the transition, requiring comparatively large tax rates for the government to equalize revenues and expenses in all periods. This constitutes a potentially large transfer from early generations to later generations.

Instead, I set a time-invariant λ for all newborn households living under the consumption tax regime. Any shortfalls in the sequential budgets are absorbed by new issues of government debt. This debt will accumulate over the transition until it settles at its new long-run level. I also keep the retirement benefit b fixed.

A related issue concerns the treatment of bequests. Because a consumption tax postpones tax liabilities until later in life, households naturally accumulate more assets over their working years. Their retirement savings must account not just for targeted consumption needs but also any tax obligations triggered by dis-saving. As a result, accidental estates are mechanically larger in the new steady state. This introduces a potential bias against newborn generations early in the transition. If the 'bequest budget' were balanced year-by-year, these generations would receive the same bequests as their labour-tax-paying neighbours. Unlike those neighbours, however, the newborn generations would owe tax when they spend their bequests. To sidestep this problem, I assign the new (initial) steady state bequest to all newborn (old) generations in each year of the transition. Any shortfall in the bequest budget is covered by the government and added to the public debt.

There are therefore only two fiscal constitutions. Let subscripts 0 and ∞ denote the initial and terminal steady states, respectively. Pre-reform households of all ages continue to live under the status quo policy $\{\tau^k, \tau^c, \tau, \lambda_0, b_0, q_0\}$ for the remainder of their lives. Post-reform households of all birth cohorts live under the new policy $\{\tau^k, \tau^c, \tau, \lambda_\infty, b_\infty, q_\infty\}$ for the entirety of their lives. These policy parameters are reported in Table.

Table 4: The Two Fiscal Constitutions

	tax base	τ	λ	τ^k	τ^c	b	q
Baseline	labour income	0.13615	0.6659	28.3%	4.4%	0.1980	0.0331
Reform	consumption	0.13615	0.6591	28.3%	4.4%	0.2456	0.0543

In summary, all households alive at time $t = 0$ continue to pay taxes on their labour income according to the same rate schedule as before. Agents who enter the economy at time $t = 1$ or later will instead pay taxes on their consumption, with the same parameter τ but with λ selected to ensure the government's intertemporal budget constraint is satisfied. Government debt adjusts along the transition to absorb any shortfalls. These policy design choices ensure that there are only two operative channels: the direct tax base effect and the associated general equilibrium effects (if any). Assuming that the economy converges to its new steady state after G periods, the induced transition path is characterized by sequences of prices $\{r_t, w_t\}_{t=1}^{G}$ and debt $\{B_t\}_{t=1}^{G}$.

7.2 Quantitative Results

7.2.1 Long-Run Impact

The considered tax reform has few consequences for existing households or the macroeconomy in the short run by construction. Only in the long run, as labour-income-taxpayers die out and consumption-taxpayers take over, do aggregate quantities reveal the impact of the policy change. A comparison of stationary equilibria is therefore sufficient to effectively demonstrate the relative merits of the two fiscal regimes. Columns 1 and 2 of Table 5 report the relevant details of the pre- and post-reform steady states. Unless otherwise indicated, interpret all quantities as per capita measures, in the tables and charts as well as in the text.

Hours and Productivity The tax reform reduces aggregate work hours by 0.1% in the long run. But because work decisions are less distorted, these hours are allocated more efficiently. Consequently, aggregate labour supply, measured in productivity units, increases by 1.1%. This improvement springs from two possible sources. First, households can work fewer hours during predictably low-wage years and more hours during predictably high-wage years. Second, households can intensify their labour supply responses to productivity shocks, in both positive and negative directions.

It turns out that intertemporal re-allocation does not play an important quantitative role. Figure 3 plots the percentage change in average hours between the two steady states over the working life (in blue). The change is slightly positive through the first two-thirds of the agent's career, but drops off substantially after wages peak at age 50 (vertical dashed line). The overall effect is marginally negative. The U-shape pattern in the first half of the working life is an artifact of the strict borrowing constraint, which binds more frequently for the very youngest households.

The household's heightened response to productivity shocks is the important channel. A system that taxes earnings period-by-period penalizes workers for aggressively exploiting temporary wage changes. This penalty disappears under a consumption tax, since the worker can reduce his tax liability by smoothing his consumption. As a result, we see significant re-allocation of effort across *states*, with workers expanding and contracting their labour supply more freely in response to positive and negative changes to their earning power. Heightened sensitivity to wage shocks leads to significantly higher average productivity in the new steady

Table 5: The Two Steady States

	(1) Baseline	(2) Reform
Tax regime		
Tax base	Earnings	Consumption
Tax level λ	0.666	0.659
Tax progressivity τ	0.136	0.136
Quantities (%Δ)		
Hours worked		-0.13
Labour supply		1.07
Capital stock		1.85
Consumption		1.49
Output		1.33
Ratios		
Investment–Output	0.250	0.252
Government–Output	0.170	0.168
Capital–Output	2.75	2.76
Debt–Output	0.97	3.22
Prices		
Wage rate	1.000	1.003
Interest rate (%)	4.00	3.94
Welfare		
CEV (%Δ)		0.92
Gini coefficients		
Consumption	0.201	0.200
Earnings	0.343	0.364
Wealth	0.495	0.506

state at every age, as shown by the green line in Figure 3.

I decompose the aggregate change in labour efficiency as follows. Let $\bar{n}_{0,j}$ and $\bar{n}_{\infty,j}$ denote mean labour supplied at age j in the initial and terminal steady states, respectively. Define $\bar{h}_{0,j}$ and $\bar{h}_{\infty,j}$ analogously for hours worked. Then I measure the net impact of the intertemporal effect by supposing that life-cycle hours conform to the terminal steady state but, counterfactually, mean productivity follows the life-cycle pattern of the initial steady state. Similarly, the net impact of the intratemporal effect imagines that hours conform to the initial steady state but productivity follows the life-cycle pattern of the terminal steady

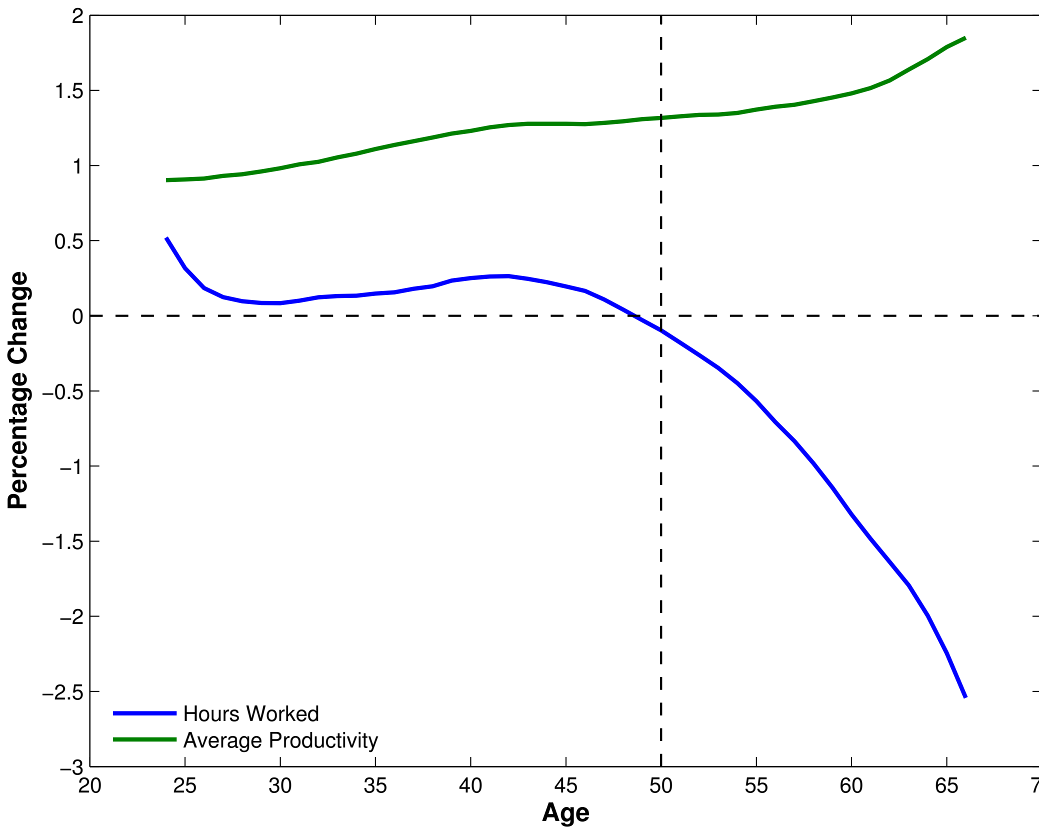

Figure 3: Steady-State Differences in Hours and Productivity Over the Life Cycle

state. That is, I calculate:

$$\text{intertemporal effect} = \sum_{j=1}^{j_R} \left(\frac{\bar{n}_{0,j} \cdot \bar{h}_{\infty,j}}{\bar{h}_{0,j}} \right) / \sum_{j=1}^{j_R} \bar{h}_{\infty,j}$$

$$\text{intratemporal effect} = \sum_{j=1}^{j_R} \left(\frac{\bar{n}_{\infty,j} \cdot \bar{h}_{0,j}}{\bar{h}_{\infty,j}} \right) / \sum_{j=1}^{j_R} \bar{h}_{0,j}$$

I report these calculations in Table 6. Note that the aggregate change in labour efficiency between steady states of 1.2% is due entirely to *intratemporal* re-allocation of hours across productivity states.

Table 6: Decomposing the Change in Aggregate Labour Efficiency

Δ% in labour efficiency due to..	Effect Size (%)	Share of Total (%)
intertemporal re-allocation of hours	-0.03	-2.3
intratemporal re-allocation of hours	1.23	102.4
combined re-allocation of hours	1.20	100.0

Labour Supply Elasticities We can further illustrate the reform's supply-side effects by recovering the elasticity of hours worked with respect to innovations to wages. Recall from (18) that we can write log-wages for household i at age j as:

$$\ln \text{wage}_{ij} = \ln w_j + \beta_{0m} + \beta_{1i}(j-1) + \beta_{2i}(j-1)^2 + \sum_{m=1}^{j-1} \phi^{j-m}\nu_{im} + \nu_{ij} + \varepsilon_{ij}$$

where ν's and ε's denote productivity shocks. Now consider a linear regression equation of the following form:

$$\ln h = \epsilon_0 + \epsilon_p \nu + \epsilon_t \varepsilon + \text{error} \tag{19}$$

The slope coefficients capture the *contemporaneous* labour supply response to persistent wage shocks ν and transitory wage shocks ε. To obtain these coefficients, I simulate a large number of histories under the labour tax regime of the initial steady state and the consumption tax regime of the final steady state. The sequences of randomly generated innovations is the same for both sets of histories. I then use the simulated data to run the regression in (19).

Pooling the data across age groups yields the aggregate elasticies reported in Table 7. We observe two things. First, and unsurprisingly, model households are more responsive to transitory shocks than they are to persistent shocks. Second, and more importantly, model households are more responsive to shocks of both kinds under the consumption tax.

Table 7: Comparing Labour Supply Elasticities Across Tax Regimes

Parameter		Labour Tax	Consumption Tax
persistent-shock elasticity	ϵ_p	0.2506	0.3092
transitory-shock elasticity	ϵ_t	0.3495	0.4294

Note: The simulated sample covers 13,500 households over 42 years of working life, for a total of 567,000 observations.

By repeating the same regression year-by-year, we can track how labour supply elasticities evolve over the life cycle. The results of this exercises are plotted in Figure 4. Households are much less responsive to wage shocks early in the life cycle when assets are low and many are borrowing-constrained. As households approach retirement, the persistent-shock elasticities (dashed lines) converge to to the transitory-shock elasticies (solid lines), as expected. As with the pooled results, the important takeaway here is that households are uniformly more sensitive to wage shocks when subject to a consumption tax.

Aggregate Quantities The tax reform also leads to higher capital accumulation, with the long-run capital stock settling close to 2% above its initial level. Combined with the increase in labour supply, this leads to a long-run increase in GDP of 1.3%. While some

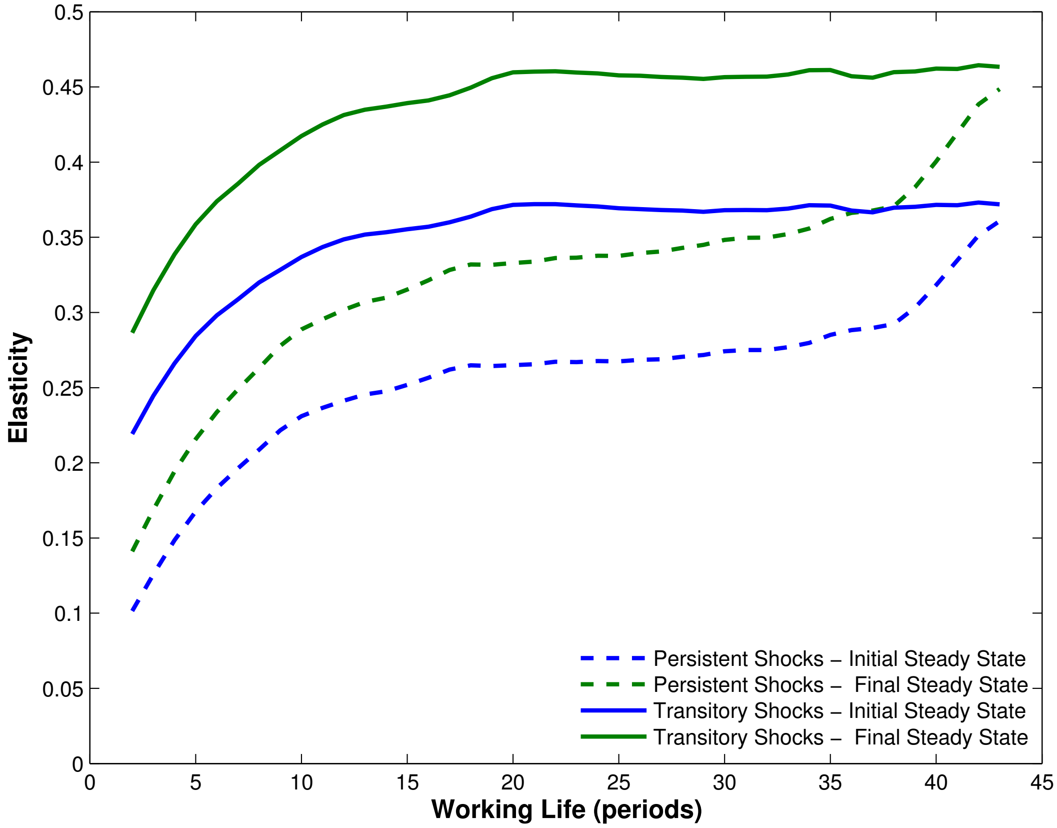

Figure 4: Labour Supply Elasticities Over the Life Cycle

of this additional output is used to maintain the larger stock of machinery and equipment, aggregate consumption still increases by a substantial 1.5%.

Distributional Effects Table 5 reports Gini coefficients for several relevant variables. Using this measure of inequality, we see that mitigating the distortion on labour effort amplifies the dispersion in labour earnings and wealth. But consumption inequality remains roughly unchanged. Thus, the increase in consumption does not come at the expense of long-run distributional concerns. Resources are better exploited but not less equally distributed in the new steady state.

Welfare Effects To measure the overall impact on welfare, I solve the following equation for the consumption equivalent variation:

$$W((1 + CEV)c_0, h_0) = W(c_\infty, h_\infty)$$

where W denotes *ex ante* expected lifetime utility (see Equation 10) and (c_0, h_0) and (c_∞, h_∞) denote allocations in the old and new steady states, respectively. In words, the CEV is the uniform percentage change in consumption at all ages and all states of the world required to make the initial allocation as attractive to a future household as the terminal allocation. Because households in the new stationary equilibrium enjoy higher consumption without working more hours, it should not surprise that the CEV is positive. The proposed reform generates a moderate long-run consumption-equivalent welfare gain of 0.9%.

Taxation Patterns Table 8 reports tax collections by source as a share of total revenue. The household tax includes retirement benefits. For comparison's sake, I also report the equivalent figures from the benchmark economy in Kindermann and Krueger (2017). The sales tax in my baseline calibration yields a slightly smaller share of revenue than in theirs. They set the sales tax rate somewhat higher at 5%. Note that capital income taxes account for a larger share of government revenue in the terminal steady state, a mechanical response to greater demand for assets in a consumption-based tax system.

Table 8: Sources of Government Revenues

	Baseline	Reform	K & K
Capital income tax	21.2%	25.7%	20.9%
Proportional sales tax	12.7%	9.7%	15.5%
Household tax-and-transfer	66.0%	64.5%	63.6%

Figure 5 illustrates the impact of the tax base on the life-cycle pattern of tax obligations. I calculate mean net taxes by age in both steady states, where net tax is defined as the sum of capital income taxes, sales taxes, and the household tax net of social security. I then discount these series with the respective interest rates to obtain present values. The top panel of Figure 5 plots the life-cycle profile of discounted net taxes for both steady states. The bottom panel plots the cumulative share of lifetime taxes paid. At the mid-career point (shown with a vertical dashed line), the average labour-income-taxpayer has paid two-thirds of her lifetime net tax burden while the average consumption-taxpayer has paid just half. It is clear from both panels that retired households in the initial steady state receive, on average, more money from the government in the form of social security benefits than they pay in capital income taxes and sales taxes. In contrast, the typical retiree in the new steady state remains a positive net contributor to the public purse well into her 80s.

7.2.2 The Transition Path

Figure 6 documents the evolution of key macroeconomic aggregates and prices along the transition path. Notice the absence of discontinuous jumps at the outset of the transition. Although sharp immediate reactions are typical with this sort of quantitative exercise, the

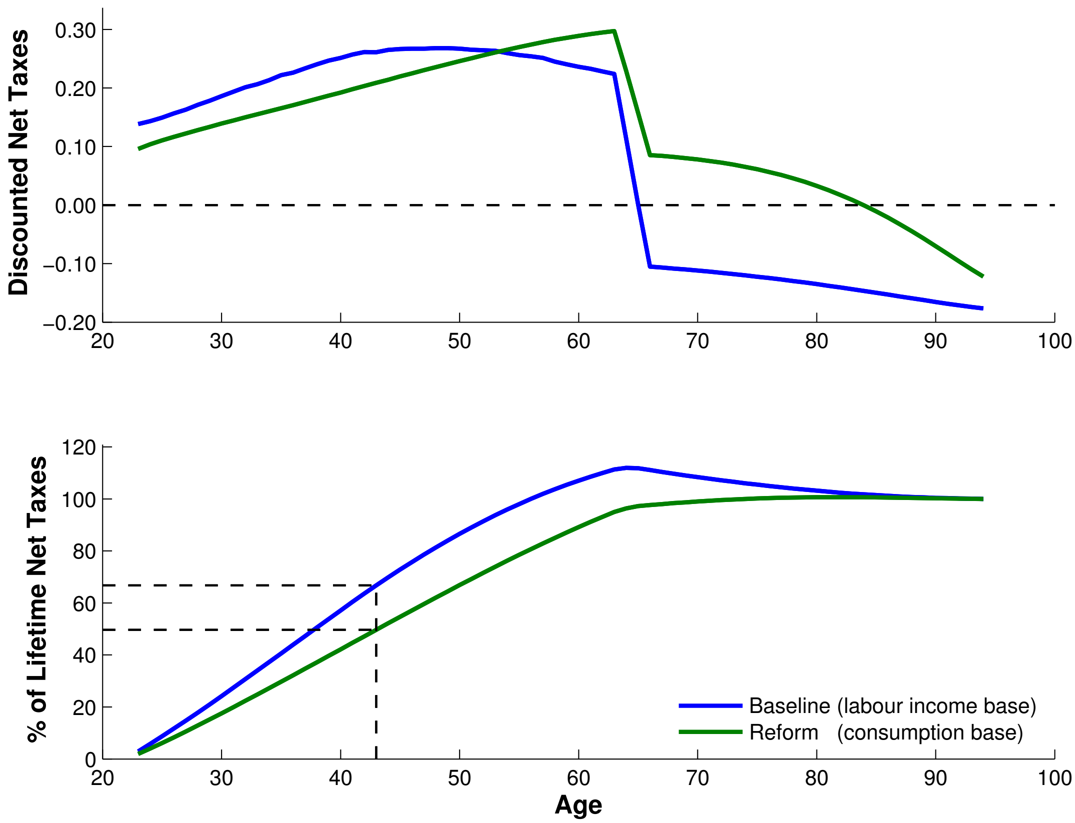

Figure 5: Discounted Net Taxes Over the Life Cycle

short-run macroeconomic response here is muted because the new policy applies only to newborn generations. Consequently, the economy evolves gradually, converging to the new steady state after one hundred periods or so.

The short-run impact on work effort is mildly positive, with average hours climbing gently over the first two decades of the transition, falling sharply thereafter to the new long-run level. Aggregate labour supply increases steadily before leveling off around the time the last of the existing generations retires from the workforce. Physical capital, total output and consumption all increase monotonically along the transition path. Most of the gains are realized by the time the first newborn generation dies out.

The reform's immediate impact on labour supply causes wages to fall and interest rates to rise in the short run, though the magnitude of these changes is negligible. After a couple of decades, the effect of higher capital accumulation kicks in and prices begin moving in the opposite directions toward new steady-state values. The real wage ultimately climbs 0.3% while the interest rate falls from 4.0% to 3.9%.

One of the most notable changes is the increase in government debt, which more than triples over the transition on a per capita basis. In the initial stationary equilibrium, gov-

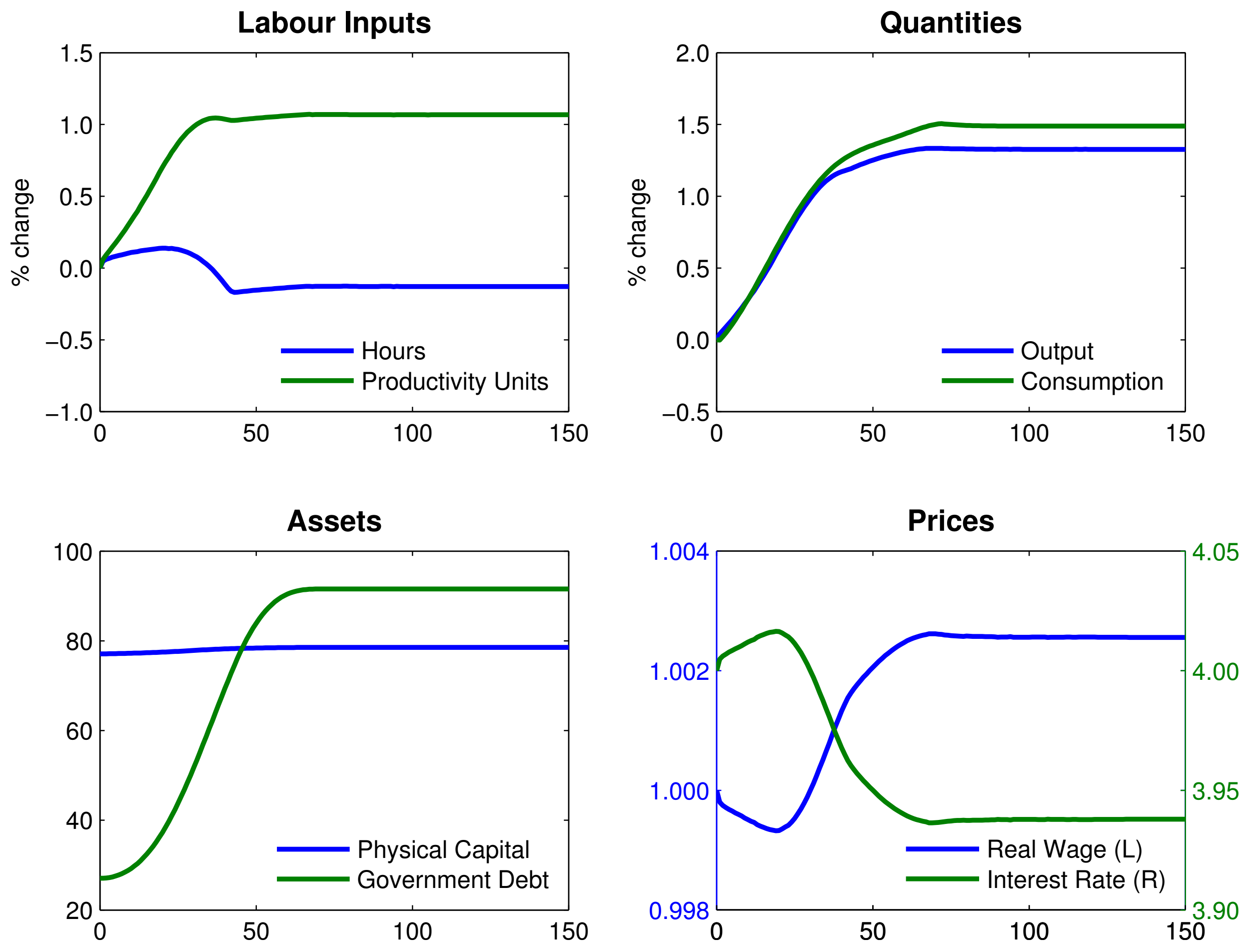

Figure 6: Transition Path for Aggregate Quantities and Prices

ernment debt accounts for roughly one-quarter of household assets, with physical capital accounting for the rest. By the time the economy converges to the post-reform steady-state, the average portfolio holds more government debt than physical capital. This massive expansion of the public debt is largely mechanical. Because tax liabilities are deferred under the consumption tax regime, the government collects fewer taxes in the early years of the transition. The government offsets these revenue shortfalls by issuing new bonds. Fortunately, there are many willing buyers for this debt since newborn cohorts know they must eventually finance the significant tax liabilities that their late-in-life consumption will trigger. Thus, the net flow of funds between households and the government is, to a rough approximation, no different under the new tax regime. The only difference is that collections are tilted more toward loans and less toward taxes. Indeed, 99% of the increase in household saving consists of government bond purchases.

The rapid and substantial increase in government debt is therefore an artifact of changes to the timing of tax collections over the life cycle, not a sign of weak government finances.

That being said, this accumulation of debt underlines the importance of computing the entire transition path when evaluating a policy change. Otherwise, a considerable part of the future fiscal burden will be covertly shifted to the near term, yielding very misleading estimates of the reform's long-run effects. My approach avoids this bias by explicitly ensuring equal tax treatment for all of the transition's cohorts. Consumption-taxpayers face the same rate schedule no matter when they are born.

As a check on how well my approach isolates the impact of the choice of tax base, we can chart welfare gains and losses by cohort. The top panel of Figure 7 displays cohorts already alive at the time of the reform. The bottom panel displays cohorts born into the transition. Recall that all post-reform generations are subject to the exact same fiscal policy; the only difference in their economic environments is the path of equilibrium factor prices which evolve over the transition as shown in Figure 6. The same can be said of all pre-reform generations. Thus, we should expect only minimal variation in welfare consequences within each panel. And this is indeed what we see.

The CEVs for existing generations are tiny and range from -0.06% for the very youngest to 0.02% for middle-aged households. The CEVs for newborn generations range from 0.85% for households born in the 17^{th} year of the transition to roughly 0.92% for all cohorts born after the first fifty years. These slight variations in welfare across cohorts are due to small changes in the evolution of wages and interest rates from one steady state to another.

7.2.3 Decomposing the Macroeconomic Effects

The reform's long-run effects stem from several operative channels. The direct effects include efficiency gains from mitigating the distortions on work decisions and possible impacts on the social insurance system. Then there are the general equilibrium prices effects. The larger capital stock in the new steady state leads to (slightly) higher wages and (slightly) lower returns to saving. I assess the relative importance of these channels by computing the transitions implied by a series of appropriate counterfactual conditions.

In the first of these exercises, the results of which are reported in column 2 of Table 9, I isolate the impact of the labour supply distortions. In this scenario, no actual changes are made to the tax base. All cohorts, old and new alike, continue to pay taxes on the basis of their earnings. Newborn households, however, are assumed to act *as though* they were subject to a newly introduced consumption tax. Their decision rules solve an auxiliary problem, one where taxes are levied on consumption, not the actual problem where taxes are levied on earnings. These households are, in a sense, 'tricked' to behave in a manner that produces the efficiency gains associated with the switch to a consumption base, but without introducing real changes to the structure of social insurance. In addition, factors prices are

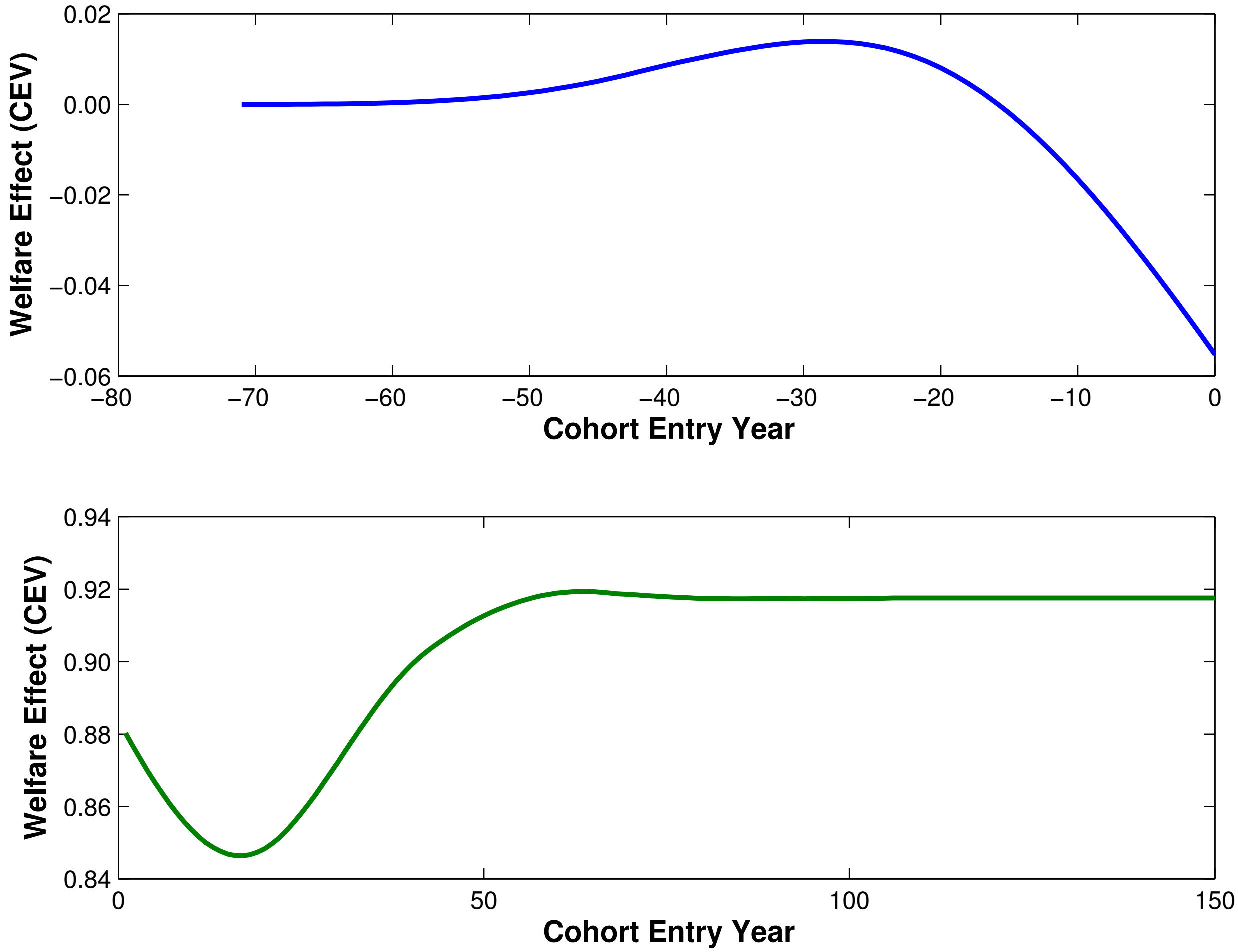

Figure 7: Welfare Effects of Policy Reform by Age Cohort

fixed at their initial steady-state levels.

The third column reports the results of the reverse exercise: households continue acting as though they are subject to tax on earnings, but tax burdens are in fact assessed according to expenditures. The conceit of this exercise is to isolate the impact on the social insurance system. Does a consumption-based tax code do a better job of redistributing resources from the rich and lucky to the the poor and unlucky? As with the first counterfactual exercise, prices are held fixed.[29]

I then quantify the joint impact of the efficiency and insurance channels by solving for the terminal steady state in the case of an open economy. The only difference between this scenario and the full reform is that prices do not evolve to reflect changes in aggregate quantities. The results of this exercise are reported in column 4 of Table 9. In all decomposition

[29]For additional details on the computation of the pseudo-consumption tax case and the pseudo-labour income tax case, see appendices C.4 and C.5.

Table 9: Decomposing the Macroeconomic and Welfare Effects

	(1) Baseline	(2) Efficiency	(3) Insurance	(4) Fixed Prices	(5) Reform
Tax regime					
Tax base	Earnings	Earnings	Consumption	Consumption	Consumption
Tax level λ	0.666	0.674	0.653	0.659	0.659
Tax progressivity τ	0.136	0.136	0.136	0.136	0.136
Quantities (%Δ)					
Hours worked		0.35	-0.41	-0.17	-0.13
Labour supply		1.59	-0.42	1.03	1.07
Capital stock					1.85
Consumption		2.31	-0.57	1.56	1.49
Output					1.33
Prices					
Wage rate	1.000	1.000	1.000	1.000	1.003
Interest rate (%)	4.00	4.00	4.00	4.00	3.94
Welfare					
CEV (%Δ)		1.35	-0.33	0.95	0.92
Gini coefficients					
Consumption	0.201	0.202	0.201	0.203	0.200
Earnings	0.343	0.364	0.352	0.364	0.364
Wealth	0.495	0.496	0.503	0.505	0.506

exercises, it is necessary to compute the transition path of pubic debt in order to obtain the correct budget-balancing values for tax parameter λ. The true initial and terminal steady states are displayed in columns 1 and 5 of Table 9.

Several observations from these exercises are worth making. First, the quantitatively most important effect is the mitigation of labour supply distortions. Eliminating these distortions, in isolation, produces a 0.4% increase in labour supply and a 1.2% improvement in labour efficiency. Aggregate consumption rises by a sizable 2.3% in the long run, more than enough to offset somewhat higher work hours. This channel is responsible for generating all the positive welfare gains associated with the simple reform.

Introducing the social insurance effect attenuates or reverses many of these aggregate impacts. For example, long-run hours are 0.4% higher after mitigating the labour supply distortion, but 0.4% lower after changing the basis for social insurance. There is no impact on aggregate labour productivity. From a welfare perspective, the social insurance effect is modestly negative. In this setting, the advantage of taxing the superior signal for lifetime resources is clearly dominated by whatever advantage lies in taxing earnings year by year.

Given the small differences in factor prices from one steady state to the other, it is not surprising to see that the small open economy departs only marginally from the closed economy. Equilibrium price effects are simply not large enough to colour the welfare consequences of the tax reform.

8 An Optimal Tax Reform

In the previous section I assessed the impact of converting the tax system from an earnings base to an expenditure base while holding progressivity fixed. This simple reform generates a sustained consumption-equivalent welfare gain of roughly 0.9%for cohorts born into the transition.

However, there is no reason to think *a priori* that the baseline progressivity is optimal for either tax base. The parameter τ reflects the tax code as it is, not as it ought to be. The results of the previous section are therefore an approximation of the differences between a progressive labour income tax and a progressive consumption tax, not its definitive measure. To obtain the latter, we must numerically characterize the *optimal* choice of τ under *both* tax regimes and then compare. The differences that we observe from the best-on-best comparison could be larger or smaller than before.

To ensure a consistent analysis, the optimal tax reform proceeds along the exact same lines as the experiment described in subsection 7.1. In particular, only newborn cohorts are subject to the new tax code and older cohorts continue to operate under the old policy rules. This is true for all experiments, regardless of whether the tax base is converted to consumption or not.

8.1 The Optimal Tax Codes

Figure 8 plots welfare gains against the progressivity parameter τ. The green and blue dashed lines indicate the locations of the optimal progressivity parameter for the labour income tax and consumption tax, respectively. The vertical black dashed line indicates the location of the baseline progressivity parameter.

There are several important things to notice from this picture. First, the optimal tax code is less progressive, in the sense that τ is lower, whether or not the tax system is converted to a consumption base. Second, the τ that maximizes welfare for the consumption-based system is higher than the one that maximizes welfare for the earnings-based system. Third,

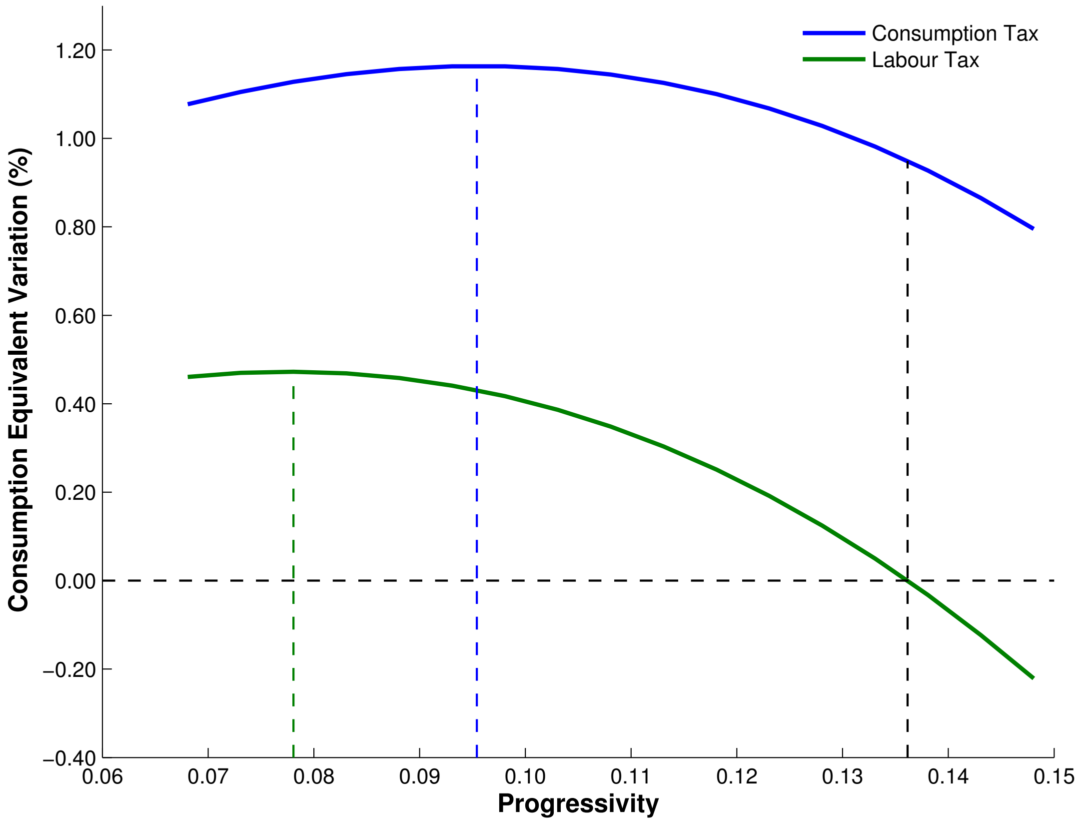

Figure 8: Long-Run Welfare Gains by Tax Base as a Function of τ

the utilitarian gap between the two tax bases narrows slightly when we compare optima.

Table 10 documents important long-run differences across tax regimes. The numbers for Baseline and Simple Reform are reproduced from Table 5. To those I have added model data for the steady states induced by a transition to the optimal labour income tax (column 2) and the optimal consumption tax (column 4). Both optimal reforms, which reduce the scale of the progressivity parameter, lead to large long-run increases in aggregate quantities. Labour supply and physical capital expand by roughly 3% and 7%, respectively, and both output and consumption increase well over 4%. Interestingly, the changes in these quantities are very similar under either optimized reform, despite the different tax bases and the different settings for the tax parameters.

Despite the substanial increases in per capita quantities, the long-run welfare impacts of optimized tax reform are surprisingly muted. The CEV is 0.5% for the optimal labour income tax. It is 1.2% for the optimal consumption tax, not much higher than the gain generated by base-conversion alone. Relative to the simple reform, the change in consumption is 3.12 percentage points greater under the optimal consumption tax. But the change in hours is 2.43 percentage points greater too. The disutility of the incremental work effort erodes most

Table 10: Terminal Steady States Under Optimal Tax Reform

	(1) Baseline	(2) Optimal L-Tax	(3) Simple Reform	(4) Optimal C-Tax
Tax regime				
Tax base	Earnings	Earnings	Consumption	Consumption
Tax level λ	0.666	0.712	0.659	0.699
Tax progressivity τ	0.136	0.078	0.136	0.095
Quantities (%Δ)				
Hours worked		2.48	-0.13	2.30
Labour supply		2.94	1.07	3.19
Capital stock		7.16	1.85	6.93
Consumption		4.35	1.49	4.61
Output		4.32	1.33	4.41
Prices				
Wage rate	1.000	1.013	1.003	1.009
Interest rate (%)	4.00	3.68	3.94	3.78
Welfare				
CEV (%Δ)		0.47	0.92	1.16
Gini coefficients				
Consumption	0.201	0.205	0.200	0.201
Earnings	0.343	0.358	0.364	0.352
Wealth	0.495	0.514	0.506	0.503

of the welfare improvement from higher consumption.

The simple reform produces a welfare gain of 0.9%. When comparing optima, the performance advantage of consumption taxation narrows to about 0.7%, somewhat smaller than before. In fact, starting from the initial steady state, there are more welfare gains to be had from converting to a cash-flow tax than there are from optimizing the existing labour income tax. The main result stands: adopting a progressive consumption tax generates moderate welfare gains relative to a tax on earnings.

8.2 Labour Supply in the New Steady State

In the previous section, I argued that that labour efficiency was the main channel through which the simple reform affected aggregate quantities and welfare. I conduct a similar analysis here for the optimal tax reforms.

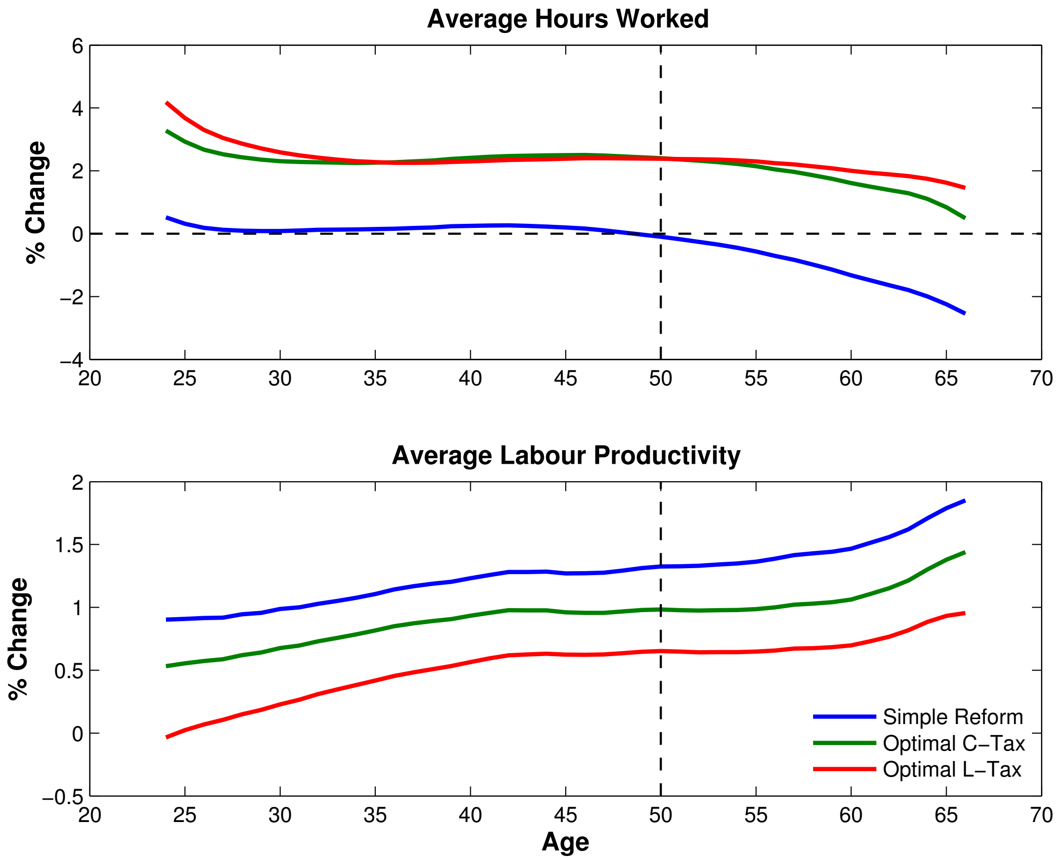

Figure 9: Hours and Productivity Over the Life Cycle by Tax Reform

In Figure 9, I trace out the life-cycle profiles of work effort and average productivity in the stationary equilibria associated with each of the three tax reforms. As shown in the top panel, the two optimal reforms induce similar changes in hours over the life cycle. The slight difference in aggregate hours between these two regimes has its source at the beginning and end of the working life. In contrast, the simple reform has a very marginally negative impact on hours. These results illustrate the impact of the progressivity parameter τ on *lifetime* hours, which applies regardless of the tax base.

The bottom panel reveals a starker rank-order with respect to changes in labour productivity, which is defined as total productivity units divided by total hours, as before. The simple reform induces the the greatest response along this dimension, greater even than the optimal consumption tax. Under the latter regime, hours are uniformly higher in all states of the world, so the change in average labour productivity is mechanically less pronounced. The optimal labour incomet tax, which continues to distort the allocation of labour across productivity states–albeit less intensely than beforee–induces the smallest improvement in labour efficiency.

Table 11 reports the steady-state changes in labour productivity, aggregated across age groups. The story for the optimal reforms is similar to that for the simple reform. There is little re-allocation of work effort across age; the improvement springs entirely from re-allocation of work effort across productivity states.

Table 11: Aggregate Labour Productivity in the New Steady State by Tax Reform

Δ% in labour efficiency due to..	Simple Reform	Optimal C-Tax	Optimal L-Tax
intertemporal re-allocation of hours	-0.03	-0.02	-0.06
intratemporal re-allocation of hours	1.23	0.89	0.51
combined re-allocation of hours	1.20	0.87	0.45

As I did for the simple reform, I estimate the elasticity of labour supply with respect to persistent and transitory innovations to the productivity process. The life-cycle elasticity profiles are plotted in Figure 10 and the aggregate elasticities are reported in Table 12. All three tax reforms induce marked increases in labour supply sensitivity to shocks of either type. The elasticty profiles of the two consumption-tax steady states are indistinguishable, while workers toiling under the optimized labour income tax are somewhat less responsive, though still much more so than in the benchmark economy.

Table 12: Comparing Labour Supply Elasticities Across Steady States

Parameter		Baseline	Simple Reform	Optimal L-Tax	Optimal C-Tax
persistent-shock elasticity	ϵ_p	0.2506	0.3170	0.2926	0.3092
transitory-shock elasticity	ϵ_t	0.3495	0.4343	0.4132	0.4308

9 Sensitivity Analysis

In this section I discuss the sensitivity of my results to the key parametric assumptions. For each alternative specification it is necessary to recalibrate the model to make the results comparable with those reported above for the baseline economy. Throughout this section I focus on results from the simple tax reform described in Section 7.

9.1 Labour Supply Elasticity

I showed in Section 7 that the welfare gains associated with a progressive consumption tax are closely linked to the willingness and ability of households to subsitute work hours across productivity states. To check if these results are quantitatively robust, I conduct a sensitivity

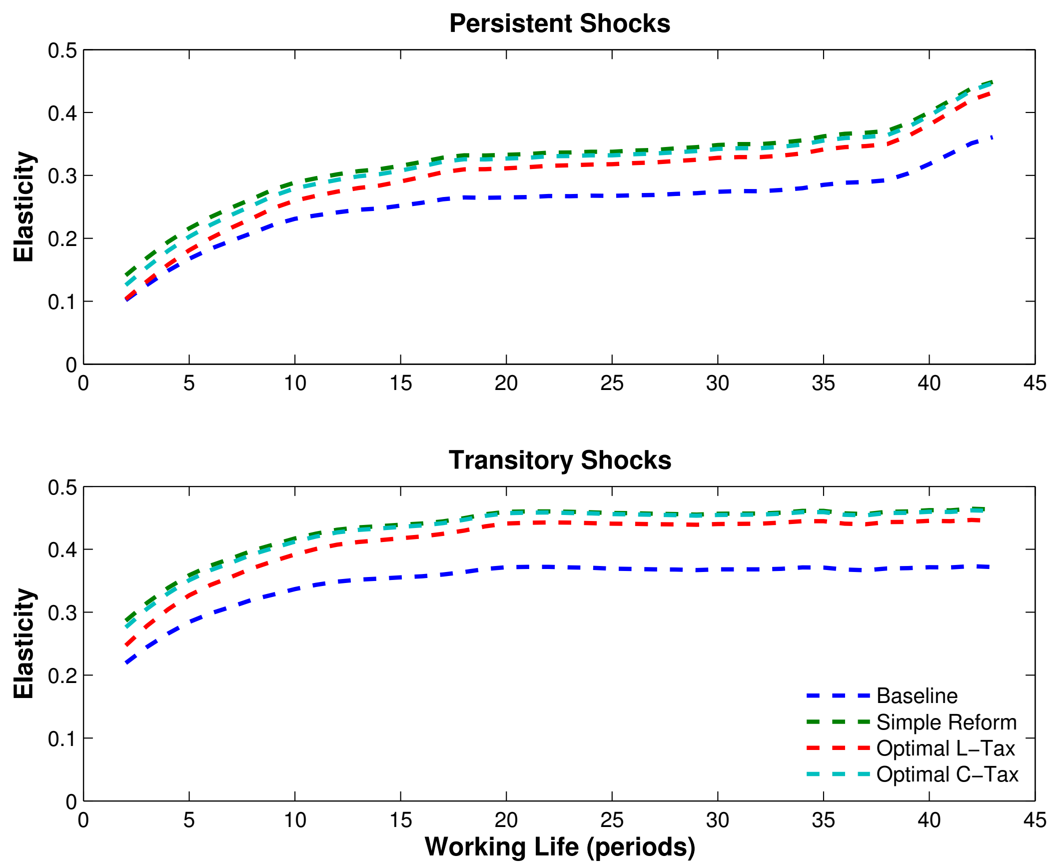

Figure 10: Labour Supply Elasticities Over the Life Cycle

analysis with respect to the labour elasticity parameter γ. In particular, Table 13 documents how the results change when we set the Frisch elasticity at a lower value (0.25) and a higher value (0.75). Not surprisingly, the impact of the simple tax reform is less dramatic when labour supply is less elastic. Long-run labour productivity improves by only 0.6% in the low Frisch scenario, about half the baseline increase. Consequently, the welfare gains are more modest, about 0.5% in consumption-equivalent terms.

9.2 Consumption Smoothing

The relative merits of progressive consumption taxation are closely linked to the household's inclination to smooth its consumer spending over time. To test how my choice for the intertemporal elasticity of substitution affects the results, I consider a lower value ($\sigma = 1$, that is, log-utility) and a higher value ($\sigma = 4$) for this parameter. The results are reported in Table 13 alongside those from the previous subsection. The key finding here is that the welfare and macroeconomic effects of the simple reform are quite sensitive to the choice of this parameter. In the log-utility case, the welfare gain of tax-base conversion is much

Table 13: Sensitivity with Respect to Preference Parameters

		Baseline	Low Frisch	High Frisch	Low CRRA	High CRRA
Parameters						
Coefficient of RRA	σ	2.00			1.00	4.00
Frisch elasticity	γ^{-1}	0.50	0.25	0.75		
Quantities (%Δ)						
Hours worked		-0.13	0.13	-0.52	-1.18	0.61
Labour supply		1.07	0.74	1.26	0.14	1.72
Capital stock		1.85	0.43	3.07	-0.75	4.90
Consumption		1.49	0.91	1.87	0.06	2.65
Output		1.33	0.63	1.85	-0.15	2.76
Prices						
Wage rate		1.003	0.999	1.006	0.997	1.010
Interest rate (%)		3.94	4.02	3.86	4.07	3.76
Welfare						
CEV (%Δ)		0.92	0.55	1.20	0.26	1.69

This table reports the long-run impact of the simple tax reform under alternative specifications of the model. The model is always calibrated so that prices in the initial steady state are $w = 1$ and $r = 4\%$, and so that one-third of the time endowment is devoted to working. The baseline results from Section 7 are reproduced her for ease of comparison. Blank cells in the parameter rows indicate no change from the baseline.

smaller, about 0.3%. In the other case, the welfare gain nearly doubles to 1.7%.

10 Conclusion

This paper asks a simple question: are consumption taxes better than labour income taxes? Departing from existing analyses in the literature, I focus attention on economies with non-linear taxes and heterogeneous agents. Equivalency between the two candidate tax systems breaks down in such settings for two reasons. First, consumption is endogenously smoothed over the life cycle and therefore serves as a less noisy signal of lifetime resources. Second, a progressive labour income tax dampens the household's responsiveness to wage changes in ways that a progressive consumption tax does not. It is the second channel that proves quantitatively important. A simple conversion of the existing tax-and-transfer system from an earnings base to an expenditure base yields a welfare gain of 0.9%, all of which is due to the re-allocation of work hours across productivity states. This relative utilitarian advantage narrows to 0.7% when we compare optimized tax systems, as shown in Section 8.

The key driver of my results is the interaction between non-linear tax schedules and fluctuating wages. But this is not the only reason to suspect the non-equivalency of consumption taxation and labour income taxation. Several others come to mind. For example, agents in my model trade a single risk-free bond. But the risk-free rate is just one possible component of the return on an investment. Investment returns also reflect risk premia, economic rents and sheer luck. An intriguing feature of consumption-based tax systems is that they effectively tax these other components while leaving the risk-free component untouched. Since most of the theoretical objections to capital income taxation concern the impact on the risk-free rate, a progressive consumption tax could allow policy-makers to tax supernormal returns (at graduated rates) without distorting the basic incentive to save and invest. This aspect of cash-flow taxation presents a promising avenue for future research.

A closely related adminstrative issue concerns the treatment of entrepreneurial income. If capital income is taxed less severely at the margin, a business owner or entrepreneur is incentivized to take dividends rather than salary, even if the income 'ought' to be thought of as labour earnings. Many problems of this sort are alleviated under a consumption tax regime.

As noted in the introduction, the cash-flow tax I study bears a resemblance to existing retirement savings programs likes IRAs and 401(k)s. The major plank of all these policies is the deductibility of contributions to savings accounts from the taxpayer's taxable income. One significant difference is that IRAs and 401(k)s are subject to many restrictions, limits and penalties. In contrast, the cash-flow tax studied here has no restraints. An important quantitative question that I leave for future research is the extent to which the restricted extant programs are able to capture the gains identified in this study.

In my model, consumption is especially smooth over the life cycle because there are no lumpy expenditures or shocks to marginal utility. If expenditures were more volatitle, the advantage of the consumption base would begin to dissipate. As discussed in Section 4, a feasible solution to such problems is to give households the choice of whether to deduct net savings on their tax returns. A hybrid system of this sort would allow households to continue saving in qualified accounts for self-insurance and retirement. But they could also use non-qualified accounts to save for durables and to weather unexpected expenditure shocks. Consequently, I do not consider the absence of such features in my model to be limiting. The main result still applies. By relaxing the distortions on household labour supply, a progressive consumption tax generates aggregate improvements in labour efficiency and non-trivial welfare gains over the long run.

References

Abowd, John and David Card, "On the covariance structure of earnings and hours changes," *Econometrica*, 1989, *57* (2), 1986–1998.

Ahsan, Syed M., "Choice of Tax Base Under Uncertainty: Consumption or Income," *Journal of Public Economics*, 1989, *40* (1), 99–134.

_ **and Peter Tsigaris**, "The Design of a Consumption Tax Under Capital Risk," *Journal of Economics*, 1998, *68* (1), 53–78.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters, and Jan Walliser, "Simulating Fundamental Tax Reform in the United States," *American Economic Review*, 2001, *91* (3), 574–595.

Andrews, William D., "A Consumption-Type or Cash Flow Personal Income Tax," *Harvard Law Review*, 1974, *87* (6), 1113–1188.

Atkeson, Andrew, V. V. Chari, and Patrick J. Kehoe, "Taxing Capital Income: A Bad Idea," *Reserve Bank of Minneapolis Quarterly Review*, 1999, *23* (3), 3–17.

Attanasio, Orazio and Thomas DeLeire, "The Effect of Individual Retirement Accounts on Household Consumption and National Saving," *The Economic Journal*, 2002, *112* (481), 504–538.

Auerbach, Alan J. and Kevin A. Hassett, "Introduction," in Alan J. Auerbach and Kevin A. Hassett, eds., *Toward Fundamental Tax Reform*, Washington, DC: American Enterprise Institution Press, 2005, pp. 1–10.

_ **, Laurence J. Kotlikoff, and Jonathan Skinner**, "The Efficiency Gains From Dynamic Tax Reform," *International Economic Review*, 1983, *24* (1), 81–100.

Bank, Steven A., "The Progressive Consumption Tax Revisited," *Michigan Law Review*, 2003, *101* (6), 2238–2260.

Bankman, Joseph and David A. Weisbach, "The Superiority of an Ideal Consumption Tax Over an Ideal Income Tax," *Stanford Law Review*, 2006, *58*, 1413–1456.

Barillas, Francisco and Jesús Fernández-Villaverde, "A Generalization of the Endogenous Grid Method," *Journal of Economic Dynamics and Control*, 2007, *31* (8), 2698–2712.

Benabou, Roland, "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 2000, *90* (1), 96–129.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten, "Consumption Inequality and Family Labor Supply," *American Economic Review*, 2016, *106* (2), 387–435.

Boadway, Robin, "Individual Income Taxation: Income, Consumption, or Dual?," Technical Report International Studies Program Working Paper 10-15, Andrew Young School of Policy Studies, Georgia State University May 2010.

Boar, Corina and Virgiliu Midrigan, "Efficient Redistribution," 2020. National Bureau of Economic Research Working Paper 27622.

Brüggemann, Bettina, "Higher Taxes at the Top: The Role of Entrepreneurs," 2020. Working paper.

Carroll, Robert and Alan D. Viard, *Progressive Consumption Taxation: The X Tax Revisited*, Washington, DC: American Enterprise Institution Press, 2012.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 2011, *101* (3), 471–75.

Coleman, Wilbur John, "Welfare and Optimum Dynamic Taxation of Consumption and Income," *Journal of Public Economics*, 2000, *76* (1), 1–39.

Conesa, Juan Carlos and Dirk Krueger, "On the Optimal Progressivity of the Income Tax Code," *Journal of Monetary Economics*, 2006, *53* (4), 1425–1450.

_ , Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 2009, *99* (1), 25–48.

Correia, Isabel, "Consumption Taxes and Redistribution," *American Economic Review*, 2010, *100* (4), 1673–94.

Daruich, Diego and Raquel Fernández, "Universal Basic Income: A Dynamic Assessment," 2020. National Bureau of Economic Research Working Paper 27351.

Diamond, Peter A., "Comment on New Dynamic Public Finance: A Users Guide," *NBER Macroeconomics Annual*, 2006, pp. 365–379.

Domenici, Pete, Sam Nunn, and Bob Kerrey, *S. 722: USA Tax Act of 1995,* 1995.

Domenici, Pete V., "The Unamerican Spirit of the Federal Income Tax," *Harvard Journal on Legislation*, 1994, *31*, 273.

Erosa, Andres and Martin Gervais, "Optimal Taxation in Life-Cycle Economies," *American Economic Review*, 2002, *90* (1), 96–129.

Fehr, Hans and Fabian Kindermann, "Taxing Capital Along the transition–Not a Bad Idea After All?," *Journal of Economic Dynamics and Control*, 2015, *51*, 64–77.

_ , _ , and Christian Habermann, "Tax-Favored Retirement Accounts: Are They Efficient in Increasing Savings and Growth?," *FinanzArchiv: Public Finance Analysis*, 2008, *64* (2), 171–198.

Feldstein, Martin, "The Effects of Taxation on Risk Taking," *Journal of Economic Theory*, 1969, *105* (2), 338–369.

Fisher, Irving and Herbert Fisher, *Constructive Income Taxation*, New York: Harper and Brothers, 1942.

Frank, Robert H., "Progressive Consumption Taxation as a Remedy for the U.S. Savings Shortfall," *Economists' Voice*, 2005, *2* (3).

Goldberg, Daniel S., *The Death of the Income Tax: A Progressive Consumption Tax and the Path to Fiscal Reform*, New York: Oxford University Press, 2013.

Guner, Nezih, Martin Lopez-Daneri, and Gustavo Ventura, "Heterogeneity and Government Revenues: Higher Taxes at the Top?," *Journal of Monetary Economics*, 2016, *56* (1), 69–85.

Guvenen, Fatih, Gueorgui Kambourov, Burhanettin Kuruscu, Sergio Ocampo-Diaz, and Daphne Chen, "Use It or Lose It: Efficiency Gains from Wealth Taxation," 2019. National Bureau of Economic Research Working Paper 26284.

Haan, Peter, Daniel Kemptner, and Victoria Prowse, "Insurance, Redistribution, and the Inequality of Lifetime Income," 2019. Working paper.

Hall, Robert E. and Alvin Rabushka, *The Flat Tax*, second ed., Stanford, CA: Hoover Institution Press, 2007.

Heathcote, Jonathan and Hitoshi Tsujiyama, "Optimal Income Taxation: Mirrlees Meets Ramsey," 2017. Working paper.

_ , Kjetil Storesletten, and Gianluca Violante, "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 2017, *132* (4), 1693–1754.

Hobbes, Thomas, *Leviathan*, London: Penguin Books, 1968. Edited by C.B. Macpherson. First published in 1651.

Huggett, Mark, "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969.

_ and Juan Carlos Parra, "How Well Does the U.S. Social Insurance System Provide Social Insurance?," *Journal of Political Economy*, 2010, *118* (1), 76–112.

_ , Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

Imrohoroglu, Ayse, Cagri Kumru, and Arm Nakornthab, "Revisiting Tax on Top Income," 2018. ANU Working Papers in Economics and Econometrics 660.

İmrohoroğlu, Ayşe, Selahattin İmrohoroğlu, and Douglas Joines, "The Effect of Tax-Favored Retirement Accounts on Capital Accumulation," *American Economic Review*, 1998, pp. 749–768.

Institute for Fiscal Studies, *The Structure and Reform of Direct Taxation*, London: George Allen and Unwin, 1978.

Kaldor, Nicholas, *An Expenditure Tax*, London: Unwin University Books, 1955.

Kaplan, Greg, "Inequality and the Life Cycle," *Econometrica*, 2012, *3* (3), 471–525.

Karahan, Fatih and Serdar Ozkan, "On the Persistence of Income Shocks over the Life Cycle: Evidence, Theory, and Implications," *Review of Economic Dynamics*, 2013, *16* (3), 452–476.

Kaymak, Barış and Markus Poschke, "The Macroeconomic and Distributional Effects of Progressive Wealth Taxes," 2020.

Keane, Michael P., "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 2011, *49* (4), 961–1075.

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1 Per Cent? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," 2017.

Kitao, Sagiri, "Labor-dependent capital income taxation," *Journal of Monetary Economics*, 2010, *57* (8), 959–974.

Kocherlakota, Narayana R., "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 2005, *73* (5), 1587–1621.

Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull, "Are Consumption Taxes Really Better Than Income Taxes?," *Journal of Monetary Economics*, 1996, *37* (3), 475–503.

Laibson, David I., Andrea Repetto, and Jeremy Tobacman, "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity*, 1998, *1998* (1), 91–172.

Laibson, David I, Andrea Repetto, and Jeremy Tobacman, "Estimating Discount Functions with Consumption Choices Over the Life Cycle," 2007. NBER Working Paper 13314.

Love, David, "Buffer Stock Saving in Retirement Accounts," *Journal of Monetary Economics*, 2006, *53* (7), 1473–1492.

Luduvice, Andre-Victor, "The Macroeconomic Effects of Universal Basic Income Programs," 2019. Working paper.

McCaffery, Edward J., *Fair Not Flat: How to Make the Tax System Better and Simpler*, Chicago: Chicago University Press, 2002.

_ , "A New Understanding of Tax," *Michigan Law Review*, 2005, *103*, 807.

Mill, John Stuart, *Principles of Political Economy*, ed. J. Laurence Laughlin, New York: D. Appleton and Company, 1885.

Mirrlees, James, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, *Tax by Design: The Mirrlees Review*, Oxford: Oxford University Press, 2011.

Motta, Giorgio and Raffaele Rossi, "Optimal Fiscal Policy with Consumption Taxation," *Journal of Money, Credit and Banking*, 2019, *51* (1), 139–161.

Nishiyama, Shinichi and Kent Smetters, "Consumption Taxes and Economic Efficiency with Idiosyncratic Wage Shocks," *Journal of Political Economy*, 2005, *113* (5), 1088–1115.

Poterba, James M., Steven F. Venti, and David A. Wise, "Do 401 (k) Contributions Crowd Out Other Personal Saving?," *Journal of Public Economics*, 1995, *58* (1), 1–32.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205–229.

Seidman, Laurence S., "Boost Saving with a Personal Consumption Tax," *Challenge*, 1989, *32* (6), 44–50.

Seidman, Lawrence S., *The USA Tax: A Progressive Consumption Tax*, Cambridge, MA: The MIT Press, 1997.

Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations*, The Glasgow Edition of the Works and Correspondence of Adam Smith, vol. II, edited by R. H. Campbell and A. S. Skinner. Oxford University Press, 1976.

Storesletten, Kjetil, Christopher Telmer, and Amir Yaron, "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 2004, *51* (3), 609–633.

Summers, Lawrence H., "Capital Taxation and Accumulation in a Life Cycle Growth Model," *American Economic Review*, 1981, *71* (4), 533–544.

_ , "An Equity Case for Consumption Taxation," in Charles E. Walker and Mark A. Bloomfield, eds., *New Directions in Federal Tax Policy For The 1980s*, Cambridge, MA: Ballinger, 1984, pp. 257–260.

United States Treasury, *Blueprints for Basic Tax Reform* 1977.

_ , *Blueprints for Basic Tax Reform*, second ed. 1984.

Vickrey, William, "Averaging of Income for Income-Tax Purposes," *Journal of Political Economy*, 1939, *47* (3), 379–397.

_ , *Agenda for Progressive Taxation*, New York: Ronald Press Company, 1947.

_ , "Tax Simplification Through Cumulative Averaging," *Law and Contemporary Problems*, 1969, *34* (4), 736–750.

— , "An Updated Agenda for Progressive Taxation," *The American Economic Review*, 1992, *82* (2), 257–262.

Weinzierl, Matthew, "The Surprising Power of Age-Dependent Taxes," *Review of Economic Studies*, 2011, *78* (4), 1490–1518.

Werning, Ivan, "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 2007, *122* (3), 925967.

Wu, Chunzan, "More Unequal Income but Less Progressive Taxation: Economics or Politics?," 2017. Working paper.

— **and Dirk Krueger**, "Consumption Insurance Against Wage Risk: Family Labor Supply and Optimal Progressive Income Taxation," 2019. Centre for Economic Policy Research Discussion Paper DP14108.

Zodrow, George R., "Taxation, Uncertainty and the Choice of a Consumption Tax Base," *Journal of Public Economics*, 1995, *58* (2), 257–265.

— , "'Should Capital Income Be Subject to Consumption-Based Taxation?," Technical Report, James A. Baker III Institute for Public Policy 2006.

A Proofs

A.1 Proof of Lemma 2

The proof for parts 1 and 2 was provided in the main text so it remains to prove part 3, which is re-stated here for convenience:

Lemma (Part 3). *Lifetime output is higher under the consumption tax—despite lifetime hours being lower—thanks, of course, to the superior allocation of effort. That is,*

$$y^c = w_1 h_1^C + w_2 h_2^C > w_1 h_1^L + w_2 h_2^L = y^L$$

Proof. To show this is true, it will be convenient to define $\varepsilon = \frac{w_2}{w_1}$ and:

$$\begin{aligned} g(\hat{\tau}) &\equiv (1-\hat{\tau})(1+\eta)(\eta+\hat{\tau})^{-1} \\ \Gamma_1(\varepsilon) &\equiv \left(1+\varepsilon^{g(\hat{\tau})}\right)^{\frac{-1}{1+\eta}} \\ \Gamma_2(\varepsilon) &\equiv \left(1+\varepsilon^{-g(\hat{\tau})}\right)^{\frac{-1}{1+\eta}} \\ \Gamma(\varepsilon) &\equiv \Gamma_1(\varepsilon) + \varepsilon\Gamma_2(\varepsilon) \\ \varphi &\equiv \left[2(1-\tau)\phi^{-1}\right]^{\frac{1}{1+\eta}} \end{aligned}$$

With this new notation, we can write hours under the labour tax as $h_j^* = \varphi\Gamma_j(\varepsilon)$ and pre-tax earnings as $y = w_1\varphi\Gamma(\varepsilon)$. Now, the only difference under the consumption tax is that the parameter $\hat{\tau}$ in the g-function is zero. So to see what happens as we adjust the distortion on the hours profile while maintaining the same progressivity, simply differentiate with respect to $\hat{\tau}$.

$$\frac{\partial y}{\partial \hat{\tau}} = \varphi\left\{\left(\frac{-1}{1+\eta}\right)\left(1+\varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)}(\ln\varepsilon)\varepsilon^{g(\hat{\tau})}g'(\hat{\tau}) - \varepsilon\left(\frac{-1}{1+\eta}\right)\left(1+\varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)}(\ln\varepsilon)\varepsilon^{-g(\hat{\tau})}g'(\hat{\tau})\right\}$$

Since $g'(\hat{\tau}) = -\left[(1-\hat{\tau})(\eta+\hat{\tau})^{-2} + (\eta+\hat{\tau})^{-1} + 1\right]$, we have:

$$\frac{\partial y}{\partial \hat{\tau}} = \left(\frac{\varphi\ln\varepsilon}{1+\eta}\right)\left[\frac{(1-\hat{\tau})}{(\eta+\hat{\tau})^2} + \frac{1}{\eta+\hat{\tau}} + 1\right]\left\{\left(1+\varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)}\varepsilon^{g(\hat{\tau})} - \varepsilon\left(\frac{-1}{1+\eta}\right)\left(1+\varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)}\varepsilon^{-g(\hat{\tau})}\right\}$$

Note that $\frac{\partial y}{\partial \hat{\tau}} = 0$ when $\varepsilon = 1$. This is sensible because when a household's wages are constant, optimal hours are constant too and no allocation distortion can apply. Otherwise:

$$\frac{\partial y}{\partial \hat{\tau}} < 0 \iff G(\varepsilon)\ln\varepsilon < 0$$

where $G(\cdot)$ is short-hand for the terms in the curly brackets.

Without loss of generality, we can examine only the case of an increasing wage profile, that is, $\varepsilon > 1$. The argument for the opposite case is symmetrical. When $\varepsilon > 1$, the distortion decreases lifetime output ($\frac{\partial y}{\partial \hat{\tau}} < 0$) if and only if:

$$\begin{aligned} G(\varepsilon) < 0 &\iff \left(1 + \varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)} \varepsilon^{g(\hat{\tau})} < \varepsilon \left(\frac{-1}{1+\eta}\right) \left(1 + \varepsilon^{g(\hat{\tau})}\right)^{-\left(\frac{2+\eta}{1+\eta}\right)} \varepsilon^{-g(\hat{\tau})} \\ &\iff \varepsilon^{2g(\hat{\tau})-1} < \left(\frac{1+\varepsilon^{g(\hat{\tau})}}{1+\varepsilon^{-g(\hat{\tau})}}\right)^{\frac{2+\eta}{1+\eta}} \\ &\iff (2g(\hat{\tau}) - 1)\ln(\varepsilon) < \left(\frac{2+\eta}{1+\eta}\right) \left[\ln(1+\varepsilon^{g(\hat{\tau})}) - \ln(1+\varepsilon^{-g(\hat{\tau})})\right] \end{aligned}$$

Define:

$$\hat{G}(\varepsilon; \hat{g}) \equiv (2\hat{g} - 1)\ln \varepsilon - \left(\frac{2+\eta}{1+\eta}\right) \left[\ln(1+\varepsilon^{\hat{g}}) - \ln(1+\varepsilon^{-\hat{g}})\right]$$

Thus, we have $G(\varepsilon) < 0$ if and only $\hat{G}(\varepsilon; \hat{g}) < 0$. Differentiating the latter function with respect to $\hat{g}$:

$$\begin{aligned} \frac{\partial \hat{G}}{\partial \hat{g}} &= 2\ln \varepsilon - \left(\frac{2+\eta}{1+\eta}\right) \left[\frac{\ln \varepsilon \cdot \varepsilon^{\hat{g}}}{1+\varepsilon^{\hat{g}}} + \frac{\ln \varepsilon \cdot \varepsilon^{-\hat{g}}}{1+\varepsilon^{-\hat{g}}}\right] \\ &= \left\{2 - \left(\frac{2+\eta}{1+\eta}\right) \left[\frac{\varepsilon^{\hat{g}}}{1+\varepsilon^{\hat{g}}} + \frac{\varepsilon^{-\hat{g}}}{1+\varepsilon^{-\hat{g}}} \left(\frac{\varepsilon^{\hat{g}}}{\varepsilon^{\hat{g}}}\right)\right]\right\} \ln \varepsilon \\ &= \left\{2 - \left(\frac{2+\eta}{1+\eta}\right) \left[\frac{\varepsilon^{\hat{g}}}{1+\varepsilon^{\hat{g}}} + \frac{1}{1+\varepsilon^{\hat{g}}}\right]\right\} \ln \varepsilon \\ &= \left(\frac{\eta}{1+\eta}\right) \ln \varepsilon \end{aligned}$$

which is striclty positive if $\varepsilon > 1$, as assumed. Recalling that $g' < 0$, this implies that

$$\frac{\partial G(\varepsilon)}{\partial \hat{\tau}} = \exp\{\hat{G}(\varepsilon)\} \frac{\partial \hat{G}}{\partial \hat{g}} \frac{\partial g(\hat{\tau})}{\partial \hat{\tau}} < 0$$

We have the following result: if $\varepsilon > 1$ and $G(\varepsilon) \leq 0$ for some $\hat{\tau}$, then $G(\varepsilon) < 0$ for any $\tau > \hat{\tau}$. So consider the no-distortion case of $\hat{\tau} = 0$ where $g(0) = \frac{1+\eta}{\eta}$. Returning to an earlier inequality, we obtain:

$$G(\varepsilon) \leq 0 \iff \varepsilon^{2\left(\frac{1+\eta}{\eta}\right)-1} \leq \left(\frac{1+\varepsilon^{\frac{1+\eta}{\eta}}}{1+\varepsilon^{-\frac{1+\eta}{\eta}}}\right)^{\frac{2+\eta}{1+\eta}}$$

Raising both sides of the inequality to the power $\frac{1+\eta}{2+\eta}$ gives:

$$G(\varepsilon) \leq 0 \iff \varepsilon^{\frac{1+\eta}{\eta}} \leq \frac{1+\varepsilon^{\frac{1+\eta}{\eta}}}{1+\varepsilon^{-\frac{1+\eta}{\eta}}} \iff \varepsilon^{\frac{1+\eta}{\eta}} \leq \frac{1+\varepsilon^{\frac{1+\eta}{\eta}}}{1+\varepsilon^{-\frac{1+\eta}{\eta}}} \iff 1+\varepsilon^{\frac{1+\eta}{\eta}} \leq 1+\varepsilon^{\frac{1+\eta}{\eta}}$$

which holds with equality. Therefore, $G(\varepsilon) = 0$ when $\hat{\tau} = 0$ (again, this is sensible because there is no distortion in this case). And based on what was already demonstrated, this implies that $G(\varepsilon) < 0$ for all $\varepsilon > 1$ and $\tau > 0$. Conclude that $\frac{\partial y}{\partial \hat{\tau}} < 0$ when $\varepsilon > 1$.

As remarked earlier, the case of a decreasing wage profile can be approached with symmetrical arguments. Putting it all together, we see that lifetime output is invariant to the tax regime for households with constant wage profiles, but is otherwise higher under the consumption tax. □

A.2 More on Lemmas 3 and 4

For sake of clarity and concision, the proofs in the main text omitted certain details, mainly algebraic ones. The proofs are produced here in full.

The Expanded Proof

Lemma. *The dynamic Ramsey problem is isomorphic to the static Ramsey problem, but only when the tax system is consumption-based.*

Proof. Using Lemma 1 to replace distutility terms in (7), the value function for the household under a progressive consumption tax is given by:

$$U^C(w;\lambda,\tau) = 2\log\lambda + 2(1-\tau)\log\left(\frac{w_1 h_1^* + w_2 h_2^*}{2}\right) - 2\left(\frac{1-\tau}{1+\eta}\right) \tag{20}$$

Now use (4) and (5) to substitute for h_1^* and h_2^*.

$$\begin{aligned}
\log\left(\frac{w_1 h_1^* + w_2 h_2^*}{2}\right) &= \log\left(\frac{1}{2}\left(w_1 + w_2 \cdot \frac{h_2^*}{h_1^*}\right) h_1^*\right) \\
&= \log\left(\frac{1}{2}\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1}{\eta}}\right)\left(2(1-\tau)\phi^{-1}\left(1+\left(\frac{w_2}{w_1}\right)^{\left(\frac{1+\eta}{\eta}\right)}\right)^{-1}\right)^{\frac{1}{1+\eta}}\right) \\
&= \log\left(\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1}{\eta}}\right)\left(1+\left(\frac{w_2}{w_1}\right)^{\frac{1+\eta}{\eta}}\right)^{\frac{-1}{1+\eta}} 2^{\left(\frac{-\eta}{1+\eta}\right)}\right) \\
&\quad + \left(\frac{1}{1+\eta}\right)(\log(1-\tau) - \log(\phi))
\end{aligned} \tag{21}$$

Substituting (21) into (20) we get:

$$U^C(w;\lambda,\tau) = 2\left\{\log\lambda + (1-\tau)\log(w_C) + \left(\frac{1-\tau}{1+\eta}\right)[\log(1-\tau) - \log\phi - 1]\right\} \tag{22}$$

where w_C denotes the agent's "pseudo-static" wage:

$$w_C = \left[w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1}{\eta}}\right]\left[1 + \left(\frac{w_2}{w_1}\right)^{\frac{1+\eta}{\eta}}\right]^{\frac{-1}{1+\eta}} 2^{\left(\frac{-\eta}{1+\eta}\right)} \tag{23}$$

The **pseudo-static wage** is the constant wage that is output-equivalent to the agent's actual wage profile. That is, for given wage profile $w = (w_1, w_2)$, the associated pseudo-static wage solves:

$$w_1 h_1^* + w_2 h_2^* = 2 \cdot w_C \bar{h}$$

where we use the fact that $\bar{h}$ is the optimal labour supply when productivity is constant over the life cycle. By rearranging this equation, it becomes clear that the pseudo-static wage can also be thought of as the household's hours-adjusted average wage:

$$w_C = \left(\frac{1}{2}\cdot\frac{h_1^*}{\bar{h}}\right)w_1 + \left(\frac{1}{2}\cdot\frac{h_2^*}{\bar{h}}\right)w_2$$

It turns out that w_C is given by (23) when the utility function and tax code take the assumed functional forms. If wages are constant, i.e., $w_1 = w_2 = \tilde{w}$, then:

$$w_C = (\tilde{w} + \tilde{w})(1+1)^{\left(\frac{-1}{1+\eta}\right)} 2^{\left(\frac{-\eta}{1+\eta}\right)} = \tilde{w}\cdot 2^{\frac{(1+\eta)-1-\eta}{1+\eta}} = \tilde{w}$$

When wages are not constant, i.e., $w_1 \neq w_2$, the pseudo-static wage is greater than than $\frac{w_1+w_2}{2}$, the arithmetic mean wage.

Notice that (22) is identical to (3), the household's value function in the static version of the model, except scaled up by the number of periods and with w_C in place of w. It is also easy to demonstrate that the government's budget constraint can be expressed in terms of the distribution of pseudo-static wages, rather than the distribution of wage profiles. Consequently, we can re-write the government's dynamic Ramsey problem as:

$$\begin{aligned} \max_{\tau,\lambda} \quad & 2\left\{\log\lambda + (1-\tau)\mathbb{E}[\log w_C] + \left(\frac{1-\tau}{1+\eta}\right)[\log(1-\tau) - \log\phi - 1]\right\} \\ \text{s.t.} \quad & 2\left\{\mathbb{E}[w_C]\bar{h} - \lambda\mathbb{E}[(w_C)^{1-\tau}]\bar{h}^{1-\tau} - g\right\} = 0 \end{aligned}$$

This multi-period problem is identical to the single-period problem R1, except that w is replaced everywhere by w_C. Hence, the dynamic problem is isomorphic to the static problem, implying that a period-by-period tax levied on current consumption can replicate a lifetime tax levied on earnings.

Can the same be said when taxes are based on labour income? The household's value function in this case is given by (7).

$$U^L(w;\lambda,\tau) = 2\log\lambda + 2\log\left(\frac{(w_1h_1^*)^{1-\tau} + (w_2h_2^*)^{1-\tau}}{2}\right) - \phi\frac{h_1^{*(1+\eta)}}{1+\eta} - \phi\frac{h_2^{*(1+\eta)}}{1+\eta}$$

Let $X = ((w_1h_1^*)^{1-\tau} + (w_2h_2^*)^{1-\tau})/2$. Then, proceeding along similar steps as before, we have:

$$\begin{aligned}
X &= \frac{1}{2}\left(w_1^{1-\tau} + \left(w_2\frac{h_2^*}{h_1^*}\right)^{1-\tau}\right)(h_1^*)^{1-\tau} \\
&= \frac{1}{2}\left(w_1^{1-\tau} + \left(w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\eta+\tau}}\right)^{1-\tau}\right)\left(2(1-\tau)\phi^{-1}\left(1+\left(\frac{w_2}{w_1}\right)^{\frac{(1-\tau)(1+\eta)}{(\eta+\tau)}}\right)^{-1}\right)^{\frac{1-\tau}{1+\eta}} \\
&= 2^{-\tau}\left(w_1^{1-\tau} + \left(w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\eta+\tau}}\right)^{1-\tau}\right)\left(2^{-\eta}\left(1+\left(\frac{w_2}{w_1}\right)^{\frac{(1-\tau)(1+\eta)}{(\eta+\tau)}}\right)^{-1}\right)^{\frac{1-\tau}{1+\eta}}\left(\frac{1-\tau}{\phi}\right)^{\frac{1-\tau}{1+\eta}}
\end{aligned}$$

Multiplying and dividing the right-hand side by $\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\tau+\eta}}\right)^{1-\tau}$ yields:

$$\begin{aligned}
X = {} & 2^{-\tau}\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\tau+\eta}}\right)^{-(1-\tau)}\left(w_1^{1-\tau} + \left(w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\eta+\tau}}\right)^{1-\tau}\right) \\
& \times\left(w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\tau+\eta}}\right)^{1-\tau}\left(\left(1+\left(\frac{w_2}{w_1}\right)^{\frac{(1-\tau)(1+\eta)}{(\eta+\tau)}}\right)^{\frac{-1}{1+\eta}} 2^{\frac{-\eta}{1+\eta}}\right)^{1-\tau} \\
& \times\left(\frac{1-\tau}{\phi}\right)^{\frac{1-\tau}{1+\eta}} \\
= {} & (\Omega(w,\tau)\cdot w_L(w,\tau))^{1-\tau}\cdot\left(\frac{1-\tau}{\phi}\right)^{\frac{1-\tau}{1+\eta}}
\end{aligned}$$

where w_L is the agent's pseudo-static wage under a labour income tax regime:

$$w_L = \left[w_1 + w_2\left(\frac{w_2}{w_1}\right)^{\frac{1-\tau}{\eta+\tau}}\right]\left[1+\left(\frac{w_2}{w_1}\right)^{\frac{(1-\tau)(1+\eta)}{(\tau+\eta)}}\right]^{\frac{-1}{1+\eta}} 2^{\left(\frac{-\eta}{1+\eta}\right)} \tag{24}$$

and Ω is defined as follows:

$$\Omega = \left[\left(\frac{1}{2^\tau} \right) \left(w_1 + w_2 \left(\frac{w_2}{w_1} \right)^{\frac{1-\tau}{\eta+\tau}} \right)^{-(1-\tau)} \left(w_1^{1-\tau} + \left[w_2 \left(\frac{w_2}{w_1} \right)^{\frac{1-\tau}{\eta+\tau}} \right]^{1-\tau} \right) \right]^{\frac{1}{1-\tau}} \tag{25}$$

Thus, we can re-write the value function as:

$$U^L(w; \lambda, \tau) = 2 \left\{ \log \lambda + (1-\tau) \log (\Omega w_L) + \left(\frac{1-\tau}{1+\eta} \right) [\log(1-\tau) - \log \phi - 1] \right\} \tag{26}$$

As under the consumption tax, the value function (26) is identical to (3), except the raw wage w is replaced by the variable Ωw_L. But the Ramsey problem can not be made isomorphic to the static version because the government budget constrained cannot be expressed in the same format. In particular, the Ramsey problem under the labour income tax is:

$$\begin{aligned} \max_{\tau,\lambda} \quad & 2 \left\{ \log \lambda + (1-\tau) \mathbb{E} \log (\Omega(w,\tau) w_L) + \left(\frac{1-\tau}{1+\eta} \right) [\log(1-\tau) - \log \phi - 1] \right\} \\ \text{s.t.} \quad & 2 \left\{ \mathbb{E}[\, \boxed{} \; w_L] \bar{h} - \lambda \mathbb{E}[(\Omega(w,\tau) w_L)^{1-\tau}] \bar{h}^{1-\tau} - g \right\} = 0 \end{aligned}$$

This problem is not identical to the static analogue because Ω does not premultiply w_L everywhere. In particular, it is "missing" at the location of the red rectangle. $\square$

Remarks The proof defined several new objects, namely the pseudo-static wages w_C and w_L and the function Ω. These objects are not mere algebraic objects, but have important economic meanings.

It is not always obvious whether one wage profile is 'better' than another. The arithmetic mean wage is not a good measure since it ignores fact that agents can allocate more effort to high-wage periods and less effort to low-wage periods. The pseudo-static wage, on the other hand, explicitly accounts for optimizing behaviour, and therefore serves as a reliable measure of lifetime earnings capacity.

Because the consumption tax leaves intertemporal work decisions undistorted, the pseudo-static wage under that tax regime provides the true ranking of lifetime productivity. If agent A's w_C is higher than agent B's, then it is correct to say that A is more productive than B. This ordering is generally not preserved under a labour tax. Moreover, the intertemporal distortion generated by progressive earnings taxation implies that $w_C > w_L$ for all wage paths. This follows directly from Lemma 2. The difference between w_C and w_L, therefore, represents the adverse effect of the intertemporal distortion.

Now, w_L is premultiplied by Ω in (26). This function has the following form:

$$\Omega^p = 2^{p-1}\left(a^p + b^p\right)(a+b)^{-p} = \left(\frac{a^p + b^p}{2}\right)\left(\frac{a+b}{2}\right)^{-p} \leq \left(\frac{a+b}{2}\right)^{p}\left(\frac{a+b}{2}\right)^{-p} = 1$$

Thus, Ω is less than unity for all wage paths, holding strictly whenever wages are not constant. It reflects the insurance penalty incurred by households with volatile wages. These households are not well served by a progressive tax on labour income, since such systems will tend to overtax them relative to their economic peers with steadier wage paths.

A.3 Proof of Proposition 3

Proposition (Age Dependence)**.** *When consumption paths are non-constant, an age-dependent period-by-period consumption tax can: (1) Eliminate the intertemporal distortion on consumption; and (2) Replicate a progressive tax on lifetime earnings.*

Proof. When the household's problem is generalized to admit age-varying tax rates, we get the following (inverse) Euler equation:

$$c_2 = \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{\beta_2}{\beta_1}\right)^{1-\tau} c_1 \implies c_2 = \left(\frac{\beta_1^{\tau}\lambda}{\beta_2^{\tau}\lambda}\right)\left(\frac{\beta_2}{\beta_1}\right)^{1-\tau} c_1 \implies c_2 = \left(\frac{\beta_2}{\beta_1}\right) c_1$$

where the second equality uses (9), the proposed age-conditioned tax plan. Notice that by allowing taxes to depend on age in the right way, we can easily eliminate the intertemporal wedge in the household's Euler equation. This modification works because it 'age-adjusts' a household's annual expenditures before assessing tax liability.

To demonstrate the second part of the proposition, I follow the same strategy as in Lemma 3. Some straightforward algebra (omitted here) yields convenient expressions for optimal hours:

$$\begin{aligned} h_1^* &= \left[(1+\beta)(1-\tau)\phi^{-1}(1+A)^{-1}\right]^{\frac{1}{1+\eta}} \\ h_2^* &= \left[\left(\frac{1+\beta}{\beta}\right)(1-\tau)\phi^{-1}(1+A^{-1})^{-1}\right]^{\frac{1}{1+\eta}} \end{aligned} \tag{27}$$

where $A = \left[\beta^{-1}\left(w_1/w_2\right)^{1+\eta}\right]^{\frac{1}{\eta}}$. Letting v denote lifetime disutility of effort, we have:

$$v = \phi\left(\frac{h_1^{*(1+\eta)} + \beta h_2^{*(1+\eta)}}{1+\eta}\right) = \frac{\phi}{1+\eta}\left\{(1+\beta)\left(1-\tau\right)\phi^{-1}\left[\left(1+A\right)^{-1} + \left(1+A^{-1}\right)^{-1}\right]\right\}$$
$$= \frac{(1+\beta)(1-\tau)}{1+\eta} \tag{28}$$

Letting y denote lifetime earnings, we have:

$$y = w_1 h_1^* + w_2 h_2^* = \left(w_1 + w_2 \cdot \frac{h_2^*}{h_1^*}\right) h_1^*$$
$$= \left(w_1 + w_2 \cdot \left[\left(\frac{1}{\beta}\right)\left(\frac{w_2}{w_1}\right)\right]^{\frac{1}{\eta}}\right)\left(\frac{(1+\beta)(1-\tau)}{\phi}\left(1+\left[\left(\frac{1}{\beta}\right)\left(\frac{w_1}{w_2}\right)^{1+\eta}\right]^{\frac{1}{\eta}}\right)^{-1}\right)^{\frac{1}{1+\eta}}$$
$$= (1+\beta) w_C \bar{h} \tag{29}$$

where $\bar{h}$ is the static labour supply, defined as before, and w_C is the agent's pseudo-static wage, re-defined as:

$$w_C = \left[w_1 + w_2\left(\frac{1}{\beta}\cdot\frac{w_2}{w_1}\right)^{\frac{1}{\eta}}\right]\left[1+\left(\frac{1}{\beta}\cdot\frac{w_2}{w_1}\right)^{\frac{1+\eta}{\eta}}\right]^{\frac{-1}{1+\eta}}(1+\beta)^{\left(\frac{-\eta}{1+\eta}\right)}$$

Similarly, an agent's lifetime tax liability is given by:

$$y - \lambda_1 x_1^{1-\tau} - \lambda_2 x_2^{1-\tau} = (1+\beta) w_C \bar{h} - \lambda\left(\frac{(1+\beta) w_C \bar{h}}{1+\beta}\right)^{1-\tau} - \left(\beta^\tau \lambda\right)\left(\frac{\beta((1+\beta) w_C \bar{h})}{1+\beta}\right)^{1-\tau}$$
$$= (1+\beta)\left(w_C \bar{h} - \lambda (w_C \bar{h})^{1-\tau}\right)$$

The household's value function is:

$$U^C = \log\left(\lambda\left(\frac{w_1 h_1^* + w_2 h_2^*}{1+\beta}\right)^{1-\tau}\right) + \beta\log\left(\beta^\tau\lambda\left(\frac{\beta(w_1 h_1^* + w_2 h_2^*)}{1+\beta}\right)^{1-\tau}\right) - \phi\left(\frac{h_1^{*(1+\eta)}}{1+\eta}\right) - \beta\phi\left(\frac{h_2^{*(1+\eta)}}{1+\eta}\right)$$

Using (28) and (29), this simplifies to:

$$U^C = (1+\beta)\left\{\log\lambda + (1-\tau)\log(w_C) - \left(\frac{1-\tau}{1+\eta}\right)\left(\log(1-\tau) - \log\phi - 1\right)\right\} + \beta\log\beta$$

Notice that the value function is a positive monotonic transformation of the analogous static value function, and in fact reduces to the benchmark two-period version when $\beta = 1$. That is,

$$U^C = (1+\beta) \cdot U^1(w_C; \tau, \lambda) + \beta \log \beta$$

The Ramsey problem can therefore be written as:

$$\begin{aligned}
\max_{\tau,\lambda} \quad & (1+\beta)\left\{\log \lambda + (1-\tau)\mathbb{E}[\log w_C] + \left(\frac{1-\tau}{1+\eta}\right)[\log(1-\tau) - \log \phi - 1]\right\} - \beta \log \beta \\
\text{s.t.} \quad & (1+\beta)\left\{\mathbb{E}[w_C]\bar{h} - \lambda \mathbb{E}[(w_C)^{1-\tau}]\bar{h}^{1-\tau} - g\right\} = 0
\end{aligned}$$

This problem is isomorphic to the static problem R1. □

A.4 Proof of Proposition 4

Proposition (Hybrid Taxation). *When consumption paths are non-constant, a period-by-period hybrid tax can: (1) Eliminate the intertemporal distortion on consumption; and (2) Replicate a progressive tax on lifetime earnings.*

Proof. The household's problem can be re-written as:

$$\begin{aligned}
U^2(w) = \max_{x_1, x_2, h_1, h_2, s} \quad & \log(\lambda x_1^{1-\tau} - s) + \beta \log(\lambda x_2^{1-\tau} + s) - \phi \frac{h_1^{1+\eta}}{1+\eta} - \beta \phi \frac{h_2^{1+\eta}}{1+\eta} \\
\text{s.t.} \quad & x_1 + x_2 = w_1 h_1 + w_2 h_2
\end{aligned}$$

Letting μ denote the Lagrange multiplier on the budget constraint, the first-order conditions are:

$$\begin{aligned}
x_1: \quad & 0 = \lambda(1-\tau)x_1^{-\tau}(\lambda x_1^{1-\tau} - s)^{-1} - \mu & (30)\\
x_2: \quad & 0 = \beta\lambda(1-\tau)x_2^{-\tau}(\lambda x_2^{1-\tau} + s)^{-1} - \mu & (31)\\
h_1: \quad & 0 = -\beta^{t-1}\phi h_1^{\eta} + \mu w_1 & (32)\\
h_2: \quad & 0 = -\beta^{t-1}\phi h_2^{\eta} + \mu w_2 & (33)\\
s: \quad & 0 = -(\lambda x_1^{1-\tau} - s)^{-1} + \beta(\lambda x_2^{1-\tau} + s)^{-1} & (34)
\end{aligned}$$

From (30) and (31), we have: $\lambda x_1 - s x_1^{\tau} = \lambda x_2 - s x_2^{\tau}$. Because $x_j \geq 0$, this equality holds if and only if $x_1 = x_2 = x$. The first-order condition for unregistered savings gives: $\lambda x_2^{1-\tau} + s = \beta(\lambda x_1^{1-\tau} - s)$. Substituting $x_1 = x_2 = x$ yields:

$$s^* = -\left(\frac{1-\beta}{1+\beta}\right)\lambda x^{1-\tau}$$

Notice that $s^* = 0$ when $\beta = 1$. When the desired consumption profile is flat, the household does not need to resort to unregistered financial vehicles to smooth its tax liabilities. A full consumption base is good enough.

The consumption allocation is given by:

$$c_1^* = \left(\frac{2}{1+\beta}\right)\lambda x^{1-\tau} \qquad c_2^* = \left(\frac{2\beta}{1+\beta}\right)\lambda x^{1-\tau}$$

and so the consumption path is intertemporally undistorted: $c_2^* = \beta c_1^* \iff u_{c_1} = \beta u_{c_2}$. From (32) and (33), we see that the Euler equation for labour supply is also undistorted: $h_1 = [\beta\,(w_1/w_2)]^{\frac{1}{\eta}}\,h_2 \iff u_{h_1} = \beta u_{h_2}$. Some further algebra yields expressions for optimal hours:

$$h_1^* = \left[(1+\beta)(1-\tau)\phi^{-1}\left(1 + \left(\beta^{-1}\,(w_2/w_1)^{1+\eta}\right)^{\frac{1}{\eta}}\right)^{-1}\right]^{\frac{1}{1+\eta}}$$
$$h_2^* = \left[(1+\beta^{-1})(1-\tau)\phi^{-1}\left(1 + \left(\beta\,(w_2/w_1)^{1+\eta}\right)^{\frac{1}{\eta}}\right)^{-1}\right]^{\frac{1}{1+\eta}}$$

Not surprisingsly, labour supply under the hybrid tax code is the same as under the age-dependent code, which means that lifetime disutility of effort is also the same.

The household's value function can be written as:

$$\begin{aligned}
U^H &= \log\left(\left(\frac{2}{1+\beta}\right)\lambda x^{1-\tau}\right) + \beta\log\left(\left(\frac{2\beta}{1+\beta}\right)\lambda x^{1-\tau}\right) - v \\
&= (1+\beta)\left\{\log\lambda + (1-\tau)\log(w_C) + \left(\frac{1-\tau}{1+\eta}\right)[\log(1-\tau) - \log\phi - 1]\right\} + \beta\log\beta \\
&= (1+\beta)\cdot U^1(w_C;\tau,\lambda) + \beta\log\beta
\end{aligned}$$

where w_C is the agent's "pseudo-static" wage, defined as under the age-dependent tax regime:

$$w_C = \left[w_1 + w_2\left(\frac{1}{\beta}\cdot\frac{w_2}{w_1}\right)^{\frac{1}{\eta}}\right]\left[1 + \left(\frac{1}{\beta}\cdot\frac{w_2}{w_1}\right)^{\frac{1+\eta}{\eta}}\right]^{\frac{-1}{1+\eta}}(1+\beta)^{\left(\frac{-\eta}{1+\eta}\right)}$$

We complete the proof in the same way as before: by observing that the hybrid Ramsey problem, expressed below, is isomorphic to the static Ramsey problem.

$$\begin{aligned}
\max_{\tau,\lambda} \quad & (1+\beta)\left\{\log\lambda + (1-\tau)\mathbb{E}[\log w_C] + \left(\frac{1-\tau}{1+\eta}\right)[\log(1-\tau) - \log\phi - 1]\right\} - \beta\log\beta \\
\text{s.t.} \quad & (1+\beta)\left\{\mathbb{E}[w_C]\bar{h} - \lambda\mathbb{E}[(w_C)^{1-\tau}]\bar{h}^{1-\tau} - g\right\} = 0
\end{aligned}$$

□

B Estimation of the Productivity Process

This appendix describes the data and estimation method used to estimate the evolution of household wages, as specified in 18. The points estimates are used to parameterize the productivity process in the model.

B.1 Data

Source The data are taken from the Panel Study of Income Dynamics (PSID), a longitudinal survey of US households beginning in 1968. Interviews occured every year until 1997, but are now conducted on a biannual basis. The initial wave consisted of nearly 2,000 families drawn from a low-income oversample and nearly 3,000 families drawn from a nationally representative core sample. These original families and their members have been tracked ever since. As sample individuals move out of existing households and form new ones, say because a child moves out of her parents' home, these newborn households are added to the sample. The decision to follow 'split-offs' helps offset attrition and adds an inter-generational dimension to the data. During the 1990s, additional households were added to the panel to correct for the absence of post-1968 immigrants.

Sample Selection A total of 28,066 individuals have appeared as a household head in the PSID. I restrict attention to white males from the core sample who satisfy the following criteria in at least four, not necessarily consecutive, waves: (i) the individual is the head of his household; (ii) the individual's age is between 24 and 65; (iii) the individual participates in the labour force (i.e. not a student, not retired); (iv) reported annual hours of work are between 500 and 5000; and (v) average hourly wages fall between \$2 and \$500.[30] Individuals who never reported years of completed schooling are also dropped. The final sample include 4286 individuals and 80,987 person-year observations (an average of 19 appearances).

Definitions The variables used in the sample selection and estimation are:

[30]Nominal variables are adjusted to a 2010 basis using the Consumer Price Index as constructed by the Bureau of Labour Statistics.

Education. The PSID records years of completed schooling for household heads. The variable is topcoded at 17 for individuals with any amount of post-graduate study. It is sometimes inconsistent as an individual might be listed as having completed 12 years of school in one year only to be listed as having completed 11 years in a later survey. To deal with this, I let a person's educational attainment be the highest education level ever reported.

Age. The age variable in the PSID does not always increase by 1 from one year to the next. Patterns such as $(30, 30, 32)$ are not uncommon, probably because interviews occur at different points during the calendar year. But patterns such as $(30, 41, 32)$ also occur, suggesting other forms of measurement error. I create a consistent age variable by inferring the year of birth that is consistent with the largest number of reports. For example, if the pattern $(30, 30, 32)$ is observed for waves $(1980, 1981, 1982)$, then I assign 1950 to be the individual's year of birth.

Labour Income. The measure of labour earnings is comprehensive and includes wages, salaries, bonuses, overtime, professional fees and commissions, as well as as the labour part of farm and business income. The variable exists for both heads and wives.

Hours Worked. The PSID records annual hours worked. This variable is constructed from answers to questions about the number of hours worked per week and the number of weeks worked per year.

Hourly Wages. The average hourly wage is the ratio of labour income to hours worked. The variable is topcoded, but some missing values can be readily recovered by calculating the wage directly.

B.2 Specification

Supposes wages evolve according to:

$$\begin{aligned} \log w_{i,t} &= \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \eta_{i,t} + \varepsilon_{i,t} + \text{year effects} \\ \eta_{i,t} &= \phi\eta_{i,t-1} + \nu_{i,t} \\ \beta_i &\sim \mathcal{N}(\beta, \Sigma) \\ \varepsilon_{i,t} &\sim \mathcal{N}(0, \sigma_\varepsilon^2) \\ \nu_{i,t} &\sim \mathcal{N}(0, \sigma_\nu^2) \quad |\phi| < 1 \quad \eta_{i,-1} = 0 \end{aligned} \tag{35}$$

where i indexes the individual and t denotes (potential) experience, that is, $t = AGE - \max\{EDUC, 12\} - 6$.

It is convenient to write the person-specific components as:

$$u_{i,t} \equiv \underbrace{\left(\tilde{\beta}_{0i} + \tilde{\beta}_{1i}t + \tilde{\beta}_{2i}t^2\right)}_{\text{heterogeneous trend}} \quad + \quad \underbrace{(\eta_{i,t} + \varepsilon_{i,t})}_{\text{stochastic disturbance}}$$

where $\tilde{\beta}_i \equiv \beta_i - \beta$. The covariance structure of $u_{i,t}$ is given by:

$$\begin{aligned} Cov(u_{i,t+k}, u_{i,t}) &= \sigma_0^2 + \sigma_1^2 t(t+k) + \sigma_2^2 t^2(t+k)^2 + \sigma_{01}(2t+k) + \sigma_{02}[t^2 + (t+k)^2] \\ &\quad + \sigma_{12}[t(t+k)^2 + t^2(t+k)] + \sigma_\varepsilon^2 \mathbb{1}\{k=0\} + \phi^k Var(\eta_{i,t}) \\ Var(\eta_{i,t}) &= \phi^2 Var(\eta_{i,t-1}) + \sigma_\nu^2 = \left(\frac{1-\phi^{2(t+1)}}{1-\phi^2}\right)\sigma_\nu^2 \end{aligned} \tag{36}$$

B.3 Identification

Suppose we have a panel of N households over T periods. With sufficient cross-sectional variation in age and experience at each point in time, we can identify the parameters of (35) as follows:

1. Since $\epsilon_{i,t} \equiv \eta_{i,t} + \varepsilon_{i,t}$ has mean zero and is independent of t and t^2, we can identify the household-specific trend parameters β_i by a linear regression argument. This in turn identifies the distribution parameters β and Σ.

2. To identify the persistence parameter ϕ notice from (36) that

$$\frac{Cov(u_{i,t+2}, u_{i,t}) - [\ldots]}{Cov(u_{i,t+1}, u_{i,t}) - [\ldots]} = \frac{\phi^2 Var(\eta_{i,t})}{\phi Var(\eta_{i,t})} = \phi$$

   where $[\ldots]$ is shorthand for terms that depend only on t and elements of Σ which are already identified.

3. The variance of persistent shocks σ_ν^2 can now be identified from any unused covariance. For example:

$$\sigma_\nu^2 = \phi^{-1}[Cov(u_{i,0}, u_{i,1}) - \sigma_0^2 - \sigma_{01} - \sigma_{02}]$$

4. Finally, any unused variance can serve to identify transitory shocks. For example:

$$\sigma_\varepsilon^2 = Var(u_{i,0}) - \sigma_0^2 - \sigma_\nu^2$$

B.4 Estimation

The wage process given by (35) is a random-effects model with serially correlated errors. I conduct the following two-step procedure to obtain estimates.

1. Run the pooled OLS regression:

$$\log \mathrm{WAGE}_{it} = \beta_0 + \beta_1 \mathrm{EXPER}_{it} + \beta_2 \mathrm{EXPERSQ}_{it} + u_{it}$$

   This yields consistent estimates for the common trend parameters β. Collect the residuals $\hat{u}_{it}$.

2. Use the first-stage residuals to construct the empirical covariance matrix. This matrix has typical element

$$\hat{C}_{t,k} = N_{t,k}^{-1} \sum_i \hat{u}_{i,t} \hat{u}_{i,t+k}$$

   where $N_{t,k}^{-1}$ denotes the number of individuals observed at both date t and $t+k$. The theoretical covariance structure given by (36) suggests a non-linear regression model of the following form:

$$y = \delta X + \gamma D + \theta g(\lambda) \tag{37}$$

   where y is a vector of empirical moments and the regressors are:

$$\begin{aligned} x_1 &= 1 & x_2 &= t(t+k) \\ x_3 &= t^2(t+k)^2 & x_4 &= 2t+k \\ x_5 &= t^2 + (t+k)^2 & x_6 &= t(t+k)^2 + t^2(t+k) \\ D &= \begin{cases} 1 \text{ if } t = k \\ 0 \text{ otherwise} \end{cases} & g(\lambda) &= \lambda^k \left(\frac{1-\lambda^{2(t+1)}}{1-\lambda^2} \right) \end{aligned}$$

   The coefficients of this model correspond directly to the remaining parameters of interest. Obtaining the NLS estimates is computationally easy since (37) is linear-in-parameters for a given value of λ (that is, for a given value of the persistence parameter ϕ). We can numerically optimize with respect to λ (that is, ϕ) by performing a simple OLS regression at each iteration.

One advantage of having rich panel data is that we can extract a large number of empirical moments to use in the second stage regression. Indeed, the number of unique elements in the variance-covariance matrix increases at an approximately quadratic rate with experience.[31] A remaining practical concern is deciding which moments to exclude (if any). Somewhat arbitrarily, I cap the covariance lag at 45 years and exclude moments to which fewer than 100 individuals contributed. [32] This leaves me with 932 empirical moments. The results of the estimation are displayed in Table 1 in the main text.

C Computation of the Models

This appendix describes the algorithms and other procedures that I use to solve the model and perform tax policy experiments.

C.1 Discretizing the Wage Process

C.1.1 The Stochastic Trend

The household's stochastic trend consists of an autoregressive component and a transitory component. Since the innovations are Gaussian, I employ a method based on Gauss-Hermite quadrature to discretize both shock processes. I use seven states for the persistent component and three states for the transitory component.

Let $\{x_i(n)\}_{i=1}^{n}$ be the roots of the n^{th} order Hermite polynomial. Then the Markov chain nodes for the persistent shock are:

$$\{\sqrt{2\sigma_\nu^2/(1-\phi^2)}x_i(7)\}_{i=1}^{7}$$

and those for the transitory shock are:

$$\{\sqrt{2}\sigma_\varepsilon x_i(3)\}_{i=1}^{3}$$

C.1.2 The Deterministic Trend

Recall that the heterogeneous trend parameters (the β_i's) are jointy drawn from $\mathcal{N}(\beta, \sigma)$ in an i.i.d. fashion. I approximate this distribution with a number of 'types', chosen in such a way that each is equally likey. This is not necessary, but it is convenient since the same number of simulations can be generated for each type.

[31] Recall that $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{1}{2}(n^2 - n)$.

[32] I also ran the regressions using only moments to which at least 200 individuals contributed. The results were robust to this change.

I select three values for each trend coefficient, implying a total of 27 types. The procedure for each coefficient is as follows:

1. Let the coefficient be (conditionally) distributed as $\mathcal{N}(\mu, \sigma)$.

2. Partition the support into three intervals: $(-\infty, A), (A, B), (B, \infty)$ where A and B are chosen so that the probability mass in each interval is the same, that is, $1/3$. Since the model is Gaussian, this means that:

$$\Phi\left(\frac{A-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{B-\mu}{\sigma}\right) = \frac{1}{3} \quad \Rightarrow \quad \begin{aligned} A &= \mu + \sigma\Phi^{-1}(1/3) \\ B &= \mu - \sigma\Phi^{-1}(1/3) \end{aligned}$$

   where Φ denotes the standard normal cdf.

3. The nodes for each partition are the conditional expectations. Recall the formula for the $\mathcal{N}(\mu, 1)$ case:

$$\mathbb{E}[X|a < X < b] = \mu + \left[\frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)}\right]$$

4. The selected nodes end up being $\{\mu - 1.09\sigma, \mu, \mu + 1.09\sigma\}$ whatever the values of μ and σ.

Since β_0 and β_1 are uncorrelated (by assumption), I can select the nodes for these two coefficients using the marginal distributions. When selecting the β_2 nodes, I need to use the conditional distribution where we are conditioning on the previously chosen β_0 and β_1. This adds an additional but straightforward step, the details of which I will not belabour here.

C.2 Computing Decision Rules: Labour Income Tax

I solve for the household's decision rules by backward induction, beginning at age J and using the endogenous grid method (EGM) to iterate on the Euler equation in reverse. This approach requires very few root-finding procedures.

Household's Recursive Problem When the tax-and-transfer system is based on earnings, the household's problem is:

$$\begin{aligned} v_t(j,a,m,n) = \max_{h,c,a'} \quad & \frac{c^{1-\sigma}}{1-\sigma} - \varphi\frac{h^{1+\gamma}}{1+\gamma} + \beta\psi_{j+1}\int v_{t+1}(j+1,a',m,\tilde{n})\pi(\tilde{n}|n)d\tilde{n} \\ \text{s.t.} \quad & c + a' = \lambda(w_t\rho(j,m,n)h)^{1-\tau} + R_t a + q_t \\ & c, a' \geq 0, \;\; h \in [0,1] \end{aligned}$$

Letting $\mu_t(j,a,m,n)$ denote the Langrange multiplier, the first-order conditions are:

$$c: \quad 0 = c^{-\sigma} - \mu_t(j,a,m,n) \tag{38}$$

$$h: \quad 0 = -\varphi h^{\gamma} + \mu_t(j,a,m,n)\lambda(1-\tau)(w_t\rho(j,m,n))^{1-\tau}h^{-\tau} \tag{39}$$

$$a': \quad 0 = \beta\psi_{j+1}\int \frac{d}{da'}\left[v_{t+1}(j+1,a',m,\tilde{n})\right]\pi(\tilde{n}|n)d\tilde{n} - \mu_t(j,a,m,n) \tag{40}$$

The envelope condition is:

$$\frac{d}{da}\left[v_t(j,a,m,n)\right] = R_t\mu_t(j,a,m,n) \tag{41}$$

Endogenous Grid Method The classic approach to solving Euler equations is to fix a state (j,a,m,n) and solve *forwards* for the optimal choice a', supposing of course that we know v_{t+1}. The EGM proposes instead that we fix a *partial* state (j,m,n) and an optimal choice a', and then solve *backwards* for the initial asset position a that rationalizes the presumed choice. Algorithm 1 details how to perform this bit of 'reverse engineering' for the benchmark model. Algorithm 3 describes how to operationalize the principle to approximate the household's entire policy function.

Algorithm 1 (Unconstrained Case)**.** *Suppose we know the marginal value of wealth in the next period, namely* $\frac{d}{da}[v_{t+1}]$*. Fix a partial state* (j,m,n) *at time* t *and consider a choice* $\hat{a}'$*. Then compute:*

1. $\tilde{v} = \beta\psi_{j+1}\int \frac{d}{da'}\left[v_{t+1}(j+1,a',m,\tilde{n})\right]\pi(\tilde{n}|n)d\tilde{n}$

and, using (38)–(40):

2. $c^* = \tilde{v}^{\frac{-1}{\sigma}}$

3. $h^* = (\lambda(1-\tau)\varphi^{-1}(w_t\rho(j,m,n))^{1-\tau}\tilde{v})^{\frac{1}{\gamma+\tau}}$

4. $\hat{a} = [c^* + \hat{a}' - \lambda((w_t\rho(j,m,n))h^*)^{1-\tau} - q_t]R_t^{-1}$

These steps yield a complete set of decision rules for the 'endogenous' state $(j, \hat{a}, m, n)$*:*

$$h_t(j, \hat{a}, m, n) = h^* \tag{42}$$

$$c_t(j, \hat{a}, m, n) = c^* \tag{43}$$

$$a_{t+1}(j, \hat{a}, m, n) = \hat{a}' \tag{44}$$

A valuable feature of EGM is that by implementing Algorithm 1 for $\hat{a}' = 0$, one can precisely identify the binding threshold for the household's budget constraint. Suppose we do just that and back out $\bar{a}$ such that $a_{t+1}(j, \bar{a}, m, n) = 0$. Then we know that the household is borrowing-constrained in all states (j, a, m, n) with $a < \bar{a}$. In these states, the household solves what is essentially a static problem. Its decisions in this case can be computed by implementing Algorithm 2.

Algorithm 2 (Constrained Case)**.** *Suppose* $a_{t+1}(j, \bar{a}, m, n) = 0$*. Consider a state* (j, a, m, n) *with* $a < \bar{a}$*. The household is borrowing-constrained so we ignore (40) and combine the other two first-order conditions to get the following necessary and sufficient condition:*

$$g(h) = h^{\gamma+\tau}\hat{\Gamma}^{\sigma} - \lambda(1-\tau)\varphi^{-1}(w_t\rho(j, m, n))^{1-\tau} = 0$$

where $\hat{\Gamma} = q_t + R_t a + \lambda((w_t\rho(j, m, n))h)^{1-\tau}$*. The derivative of* g *is*

$$g'(h) = h^{\gamma+\tau}\sigma\hat{\Gamma}^{\sigma-1}(1-\tau)\lambda(w_t\rho(j, m, n))^{1-\tau}h^{-\tau} + (\gamma+\tau)h^{\gamma+\tau-1}\hat{\Gamma}^{\sigma}$$

Notice that $g(0) < 0$*,* $g(\infty) > 0$ *and* $g' > 0$*. Thus,* g *is strictly increasing and continuously differentiable with a known derivative* g' *and a single root on* $\mathbb{R}_+$*, which, not incidentally, happens to be the solution to the household's static labour supply problem. This means we can easily apply Newton's method to find the solution. If the constrained household has zero wealth (*$a = 0$*), then the numerical root-finding can be skipped entirely as the solution has a closed-form:*

$$h_t(j, 0, m, n) = \left((1-\tau)\varphi^{-1}(\lambda(w_t\rho(j, m, n))^{1-\tau})^{1-\sigma}\right)^{\frac{1}{\gamma+\tau+\sigma(1-\tau)}}$$

It can prove useful to use $h_t(j, 0, m, n)$ *to initiate Newton's method for* $a > 0$*.*

I am now in a position to describe the full EGM for solving the benchmark model.

Algorithm 3 (EGM: Labour Tax Model)**.** *Fix the model's parameters and construct the discretized versions of* $\mathcal{M}$ *and* $\mathcal{N}$*. Also discretize the state space for the continuous asset* *variable. Let $\mathcal{A} = \{a_1, \ldots, a_{\max}\}$ denote the fixed asset grid.*[33] *The objective is to solve for the optimal decisions on $\mathcal{Z} \equiv \{1, \ldots, J\} \times \mathcal{A} \times \mathcal{M} \times \mathcal{N}$. Begin with $j = J$ and set $v_{J+1} = 0$.*

1. *Set $m = 1$ and $n = 1$.*
2. *For all $a \in \mathcal{A}$, apply Algorithm 1 with $\hat{a}' = a$ as the presumed choice. Construct the endogenous asset grid $\mathcal{G} = (\hat{a}_1, \ldots, \hat{a}_{\max})$ while looping through $\mathcal{A}$ along with the associated decision rules.*
3. *Now we know the optimal decisions for states $(j, \hat{a}, m, n)$ where $\hat{a} \in \mathcal{G}$. But we want to know the optimal decisions for states (j, a, m, n) where $a \in \mathcal{A}$. To obtain the latter, use the decision rules on the endogenous grid $\mathcal{G}$ to interpolate for the decision rules on the fixed grid $\mathcal{A}$. Note that it is possible—nay, likely—that there exist $a \in A$ such that $a < \hat{a}_1$. Do not extrapolate below. These are constrained states, so apply Algorithm 2 instead.*
4. *Select a different (m, n) and repeat steps 2 and 3 until $\mathcal{M} \times \mathcal{N}$ is exhausted.*
5. *Use (38) and (41) to compute $\frac{d}{da}[v_t]$ at every grid point in $\mathcal{Z}$. Store this for the next iteration.*
6. *Go to $j = j - 1$ and repeat steps 1-5. Stop once the steps for $j = 1$ are complete.*

Remarks:

1 Algorithm 3 is very efficient since it eliminates the need for numerical root-finding except when solving for the household's constrained problem. This is possible because of the functional forms taken by preferences and taxes. For alternative parameterizations, Algorithm 1 would require numerical root-finding when computing the optimal labour supply.

2 An alternative approach is to store the decision rules on the endogenous grid, not the fixed grid, and use these to generate simulated histories. If this option is chosen, Algorithm 2 must be applied at the simulation stage whenever a simulated household is borrowing-constrained. This means potentially many more calls to a root-finding procedure, a disadvantage. The advantage is that the model-generated data would be filtered through a single interpolation step (at simulation), possibly reducing numerical error.

[33] There are three important choices for the selection of $\mathcal{A}$: (1) the number of grid points; (2) the value of the maximal grid point; and (3) the spacing of grid points. I choose 300 grids points and let the maximal grid point equal a multiple of the highest feasible earnings, high enough so that simulated assets never exceed that level. I use a double-exponential grid so that the grid is much finer at the low end where the decision rules are more likely to be highly non-linear.

3 It is not necessary to compute the value functions $\{v_t\}$, only their wealth-derivatives. Moreover, it is only necessary to keep $\frac{d}{da'}[v_{t+1}]$ in memory.

C.3 Computing Decision Rules: Consumption Tax

There is no substantive change to the EGM algorithm when $\hat{T}$ is based on consumption instead of earnings. Only the equations used in Algorithms 1 and 2 are different. These equations are derived from the optimality conditions of the household's recursive problem, which is now formulated as:

$$\begin{aligned} v_t(j,a,m,n) = \max_{h,x,a'} \quad & \frac{(\lambda x^{1-\tau})^{1-\sigma}}{1-\sigma} - \varphi \frac{h^{1+\gamma}}{1+\gamma} + \beta\psi_{j+1} \int v_{t+1}(j+1,a',m,\tilde{n})\pi(\tilde{n}|n)d\tilde{n} \\ \text{s.t.} \quad & x + a' = w_t\rho(j,m,n)h + R_t a + q_t \\ & c, x, a' \geq 0, \;\; h \in [0,1] \end{aligned}$$

Here, x denotes expenditures net of taxes. It proves convenient to formulate the problem this way, with x as a choice variable instead of c. Letting $\mu_t(j,a,m,n)$ denote the Langrange multiplier, the first-order conditions are:

$$\begin{aligned} x: \quad & 0 = (1-\tau)\lambda^{1-\sigma} x^{-(\sigma+\tau-\sigma\tau)} - \mu_t(j,a,m,n) \\ h: \quad & 0 = -\phi h^{\gamma} + w_t\rho(j,m,n)\mu_t(j,a,m,n) \\ a': \quad & 0 = \beta\psi_{j+1} \int \frac{d}{da'}\left[v_{t+1}(j+1,a',m,\tilde{n})\right]\pi(\tilde{n}|n)d\tilde{n} - \mu_t(j,a,m,n) \end{aligned}$$

These conditions imply that under the consumption tax regime, steps 2-4 of Algorithm 1 are:

2. $x^* = \left(\frac{\tilde{v}}{(1-\tau)\lambda^{1-\sigma}}\right)^{\frac{-1}{\sigma+\tau-\sigma\tau}}$
3. $h^* = (\varphi^{-1} w_t\rho(j,m,n)\tilde{v})^{\frac{1}{\gamma}}$
4. $\hat{a} = [x^* + \hat{a}' - w_t\rho(j,m,n)h^* - q_t]R_t^{-1}$

Similarly, for Algorithm 2 we now find the positive root of:

$$g(h) = h^{\gamma}(q_t + R_t a + w_t\rho(j,m,n)h)^{\sigma+\tau-\sigma\tau} - (1-\tau)\lambda^{1-\sigma}\varphi^{-1}w_t\rho(j,m,n)$$

C.4 Computing Decision Rules: Pseudo-Consumption Tax

In Subsection 9, I run two decomposition exercises to separate the efficiency and insurance effects of the tax-base conversion. In the first of these exercises, the household is assumed to act as though it is subject to a consumption tax, but its actual tax burden is assessed according to its earnings. The basic structure of the EGM algorithm is the same as before, but the formulas for computing decisions and assets are a mixture of the consumption-tax case and the labour-income-tax case.

Given $\tilde{v}$, a pseudo-consumption taxpayer chooses expenditures and hours in the same way as a genuine consumption taxpayer. Namely:

$$x^* = \left(\frac{\tilde{v}}{(1-\tau)\lambda^{1-\sigma}} \right)^{\frac{-1}{\sigma+\tau-\sigma\tau}}$$

$$h^* = \left(\varphi^{-1} w_t \rho(j,m,n) \tilde{v} \right)^{\frac{1}{\gamma}}$$

But instead of storing x^*, which reflects expenditure net of tax, we store the associated consumption level $c^* = \lambda(x^*)^{1-\tau}$. Also, we back out start-of-period assets in a way that reflects the actual tax assessment on labour income:

$$\hat{a} = [c^* + \hat{a}' - \lambda((w_t\rho(j,m,n))h^*)^{1-\tau} - q_t]R_t^{-1}$$

In writing the code for this algorithm, special attention must be paid to the computation of the marginal utilities. The household must believe it is paying tax on consumption in the next period as well as in the present one, though of course it is doing so in neither.

C.5 Computing Decision Rules: Pseudo-Labour-Income Tax

The pseudo-consumption tax isolates the impact of the tax base reform on labour efficiency. To isolate the impact on social insurance, we perform the reverse exercise. That is, make the household acts as though it is subject to a tax on earnings, but assess actual tax burdens according to consumption.

In this case, the decisions for consumption and hours are per the equations described in Algorithm 1:

$$c^* = \tilde{v}^{\frac{-1}{\sigma}}$$

$$h^* = \left(\lambda(1-\tau)\varphi^{-1}(w_t\rho(j,m,n))^{1-\tau}\tilde{v} \right)^{\frac{1}{\gamma+\tau}}$$

But we store the required expenditure $x^* = (c/\lambda)^{\frac{1}{1-\tau}}$ and recover start-of-period assets according to:

$$\hat{a} = [x^* + \hat{a}' - w_t\rho(j,m,n)h^* - q_t]R_t^{-1}$$

C.6 Computing the Initial Stationary Equilibrium

There are seven parameters and equilibrium objects needing internal calibration. Four, viz. (β, λ, b, q), can only be calibrated by simulating the model repeatedly until specified targets are jointly attained. Two, viz. (A, B), can be normalized analytically at each iteration. The final parameter, viz. φ, can be normalized numerically after the rest of the model is calibrated.

Algorithm 4 (Initial Steady State)**.** *To solve for the stationary equilibrium, iterate on the following steps:*

1. *Fix* $r = 0.04$ *and* $w = 1.00$*.*
2. **Guess** (β, λ, b, q)*.*
3. **Solve** *for decision rules* (c, h, a') *using Algorithm 3.*
4. **Simulate** *a large number of histories. Set a specific seed for the pseudo-random number generator so that the same shock histories are used at each iteration.*
5. **Aggregate** *variables across simulated histories and compute equilibrium objects. In so doing, choose* A *and* B *so that the implied wage and implied debt-to-output ratio match their targets exactly.*[34]
6. **Verify** *if the implied values of* r*,* G/Q*, and* $b/\bar{y}$ *are sufficiently close to their targets, and if the implied* q *is sufficiently close to the guess for* q*. If so, move to the next step. If not, update the guess for* (β, λ, b, q) *and go back to step 3. Here, 'sufficiently close' means that the absolute difference is less than a given tolerance level.*
7. *The final step is to calibrate* φ *to the target for mean hours. This step is essentially a normalization since all we're doing is re-scaling the economy. Adjust all the level parameters and grids appropriately, then repeatedly apply Algorithm 3 for different values of* φ *until the target is attained. Bisection works fine here as few iterations are typically needed.*

Remark: In practice, I add another loop to Algorithm 4 by repeatedly calibrating the model for increasingly stringent tolerance levels. This ensures that the solution is approached in a comparatively uniform manner from all dimensions.

[34] Let $\overline{w}$ and $\overline{B/Q}$ denote the steady-state targets. Let $\tilde{N}$ and $\tilde{A}$ denote implied effective labour and implied total assets, aggregated over simulated histories. Since labour's share of income is $wN = (1-\alpha)Q$, it is straightforward to set $B = \overline{B/Q} \cdot \frac{\bar{w}N}{1-\alpha}$. This implies a capital stock of $\tilde{K} = \tilde{A} - B$. Then, along similar lines, we obtain the technology parameter $A = \frac{\overline{w}}{1-\alpha}\left(\frac{\tilde{N}}{\tilde{K}}\right)^{\alpha}$.

C.7 Computing the Terminal Steady State

I assume that the economy transitions to a new steady state after any change to the policy environment. Certain calibrated parameters are kept fixed, namely b, B, A and G. Certain equilibrium objects must be solved for, namely w, r, λ and q.

Algorithm 5 (Terminal Steady State)**.** *Fix a terminal public debt B'. Then iterate on the following steps until convergence.*

1. **Guess** (r', w', λ', q').

2. **Solve** *for decision rules using Algorithm 3.*

3. **Simulate** *a large number of histories.*

4. **Aggregate** *variables across simulated histories and compute equilibrium objects.*

5. **Verify** *that the implied values of r', w' and q' are sufficiently close to their guesses, and that the government budget is balanced. If not, update the guess for (r', w', λ', q') and go back to step 2.*

C.8 Transition Path

Algorithm 5 assumes a particular level of public debt B' when computing the new steady state. My algorithm for computing the transition path ensures that this choice is consistent with the behavioural changes induced by the reform.

Algorithm 6 (Transition path)**.** *Suppose that the economy is in the initial steady state at time $t = 0$ after which an unexpected policy reform is announced, effective $t = 1$. We are interested in computing the transition induced by this reform. Suppose that the economy converges to a new steady state in G periods or less. Pick G sufficiently large.*

1. **Guess** *a terminal debt B'. Apply Algorithm 5 to find the terminal steady state.*

2. **Guess** *a sequence of prices $\{r_t, w_t\}_{t=1}^{G}$.*

3. **Solve** *for the decision rules for each generation $g = -(J-1), \ldots, G$, where g indexes the period in which the cohort enters the economy. That is, generation $g = 1$ is the generation that is born in the first period of the new policy regime.*

4. **Simulate** *a large number of histories for each generation.*

5. **Aggregate** *variables across simulated histories and across generations for each time period. Compute equilibrium objects.*

6. **Verify** *that the implied values of* r_t *and* w_t *are sufficiently close to their guesses at* every *point along the transition. If not, update the guess for* $\{r_t, w_t\}_{t=1}^{G}$ *and go back to step 3.*

7. *Iterate on the governement's period-by-period budget constraint to compute the implied accumulation of public debt along the transition. Verify that the resulting terminal debt is sufficiently close to the guess. If not, update the guess for* B' *and go back to step 2.*