

**NELCO**  
**NELCO Legal Scholarship Repository**

---

New York University Law and Economics Working  
Papers

New York University School of Law

---

1-9-2007

# Beyond the Pro-Consumption Tax Consensus

Daniel N. Shaviro  
NYU, shavirod@juris.law.nyu.edu

Follow this and additional works at: [http://lsr.nellco.org/nyu\\_lewp](http://lsr.nellco.org/nyu_lewp)

 Part of the [Taxation Commons](#), and the [Taxation-Federal Income Commons](#)

---

## Recommended Citation

Shaviro, Daniel N., "Beyond the Pro-Consumption Tax Consensus" (2007). *New York University Law and Economics Working Papers*. Paper 82.  
[http://lsr.nellco.org/nyu\\_lewp/82](http://lsr.nellco.org/nyu_lewp/82)

This Article is brought to you for free and open access by the New York University School of Law at NELCO Legal Scholarship Repository. It has been accepted for inclusion in New York University Law and Economics Working Papers by an authorized administrator of NELCO Legal Scholarship Repository. For more information, please contact [tracy.thompson@nellco.org](mailto:tracy.thompson@nellco.org).

## BEYOND THE PRO-CONSUMPTION TAX CONSENSUS

Daniel Shaviro\*

**Preliminary Draft**

All Rights Reserved

### I. INTRODUCTION

For many decades in United States tax policy debate, fundamental tax reform was identified primarily with adopting a comprehensive income tax base.<sup>1</sup> In the last ten or so years, it has increasingly come to denote instead replacing the income tax with a consumption tax.<sup>2</sup> This shift has been as unmistakable in the academic literature as in public political debate.<sup>3</sup> Its occurrence is important, even though in my view the prospects for fundamental reform are decidedly dim,<sup>4</sup> because ideas and ideals matter.

In academic circles, the shift reflects a new consensus (widespread if not universal) that an ideal consumption tax is more efficient than an ideal income tax and can be equally progressive.<sup>5</sup> This implies no tradeoff between the two types of tax base of the sort that would be implied either by a choice between greater progressivity and greater efficiency,<sup>6</sup> or by a choice between different kinds of efficiency.<sup>7</sup> Rather, the ideal consumption tax is now thought to dominate the ideal income tax, in that it could be equally progressive but more efficient, or equally efficient but more progressive.<sup>8</sup>

---

\* Wayne Perry Professor of Taxation, NYU Law School. I am grateful to Alan Auerbach, Lily Batchelder, Alex Raskolnikov, Jeff Strnad, and the participants in a Harvard Law School Seminar on Current Research on Taxation for helpful comments on earlier

<sup>1</sup> Cites.

<sup>2</sup> Cites.

<sup>3</sup> Cites.

<sup>4</sup> Cite my Rice article.

<sup>5</sup> Cites.

<sup>6</sup> See, e.g., Graetz..

<sup>7</sup> See, e.g., Gruber textbook..

<sup>8</sup> Cite Bankman & Weisbach.

There is something mystifying about this shift in ideals, no matter how intellectually persuasive one finds it. Consider a second possible shift in normative framework, from favoring an annual income or consumption tax to favoring one that is lifetime-based. Under a lifetime-based system, even if tax payments are remitted annually, lifetime rather than merely annual economic results determine how much one must pay. A lifetime-based system that still employs annual tax returns can be implemented through income averaging, a system that the economist William Vickrey proposed nearly seventy years ago,<sup>9</sup> and that the U.S. federal income tax featured in a much more limited fashion from 1964 through 1986.<sup>10</sup> Income averaging has received extensive, but on the whole surprisingly unfavorable, attention in the tax policy literature<sup>11</sup> – even though, as we will see, the intellectual case for it substantially overlaps with that underlying the new pro-consumption tax consensus. At the level of ideals, leaving aside differences in administrative feasibility, it makes little sense for the one idea to be accepted while the other is rejected.

This peculiar disjuncture between the ideal tax base and income averaging debates reflects the affirmative case for income averaging has not been as widely discussed or appreciated as that for consumption taxation, nor has the overlap in rationales been generally understood. In addition, while consumption taxation certainly does not lack critics,<sup>12</sup> the relevance to the tax base debate of grounds for unease or uncertainty about income averaging remains under-appreciated. These grounds relate to difficulties with three key assumptions under which, within a standard public economics

---

<sup>9</sup> Cite Vickrey 1939.

<sup>10</sup> Cite for U.S. rules.

<sup>11</sup> See Schmalbeck, Weisbach, Buchanan, Fennell-Stark. Note Liebman is more favorable, Batchelder's more targeted proposal.

<sup>12</sup> Cites.

framework, the case for both income averaging and consumption taxation would be overwhelming. The three assumptions are as follows:

1) Complete markets – Markets are complete when they cover every possible commodity and combination thereof.<sup>13</sup> In illustration, with complete labor markets one can work at one’s wage rate for any number of hours between zero and full-time.<sup>14</sup> With complete capital markets, one can hold financial positions that would pay off in every possible state of the world, thus providing effective insurance against any possible contingency.<sup>15</sup> Complete markets are necessary to the full achievement of allocative efficiency in the economy.<sup>16</sup>

2) Consistent rational choice – Under conventional economic assumptions, people have stable preferences that determine the utility they will experience in alternative states of the world, and which they consult in order to make decisions aimed at maximizing expected utility.<sup>17</sup> People therefore are assumed to engage in consistent rational choice, given their preferences, suggesting that they will make the same choice from within a given opportunity set no matter how the choices are presented or framed.

3) Limited information– For reasons that I discuss further in section II, tax equity typically is defined in welfare economics in terms of redistribution from high-ability to low-ability individuals. The aim is to transfer resources to individuals whose marginal utility of a dollar is presumed to be greater, in practice by providing insurance against ability risk, the resolution of which is assumed to have a dominant effect on one’s total

---

<sup>13</sup> Harvey S. Rosen, PUBLIC FINANCE 50 (5<sup>th</sup> ed. 1999).

<sup>14</sup> Edward J. McCaffery, Slouching Towards Equality: Gender Discrimination, Market Efficiency, and Social Change, 103 Yale L.J. 595, 619 (1993).

<sup>15</sup> Jeff Strnad, Taxing New Financial Products: A Conceptual Framework, 46 Stan. L. Rev. 569, 578 (1994)

<sup>16</sup> Rosen, *supra*, at 50.

<sup>17</sup> See, e.g., Jonathan Gruber, PUBLIC FINANCE AND PUBLIC POLICY 24 (2004); Christine Jolls, Cass R. Sunstein, and Richard Thaler, A Behavioral Approach to Law and Economics, 50 Stan. L. Rev. 1471 (1998).

and marginal utility.<sup>18</sup> Users of this framework commonly assume that ability is the only personal attribute of distributional interest, but cannot be observed directly. Moreover, they often assume that the only way to observe it indirectly is by measuring earnings. These, in turn, are understood to provide only a flawed or noisy signal of ability, as earnings also reflect people's unobserved decisions regarding labor supply.<sup>19</sup> In effect, a high-ability person who decides not to work hard masquerades as a low-ability person, in effect pretending to qualify for insurance benefits. Importantly, the process of observing ability through earnings is commonly treated as static. That is, the government is assumed not to update its information about a particular individual by observing, say, past earnings or savings decisions.<sup>20</sup>

Why do these three key assumptions provide linked support for income averaging and consumption taxation? And why do problems with the assumptions similarly weaken the case for each? The discussion in the remainder of this article proceeds as follows. Section II further sets the context by offering a fuller description of the optimal income tax framework that supports viewing an earnings-based tax as providing ability insurance. Section III discusses the case for income averaging, which depends on the merits of an important application of the complete markets and consistent rational choice assumptions. This is the permanent income hypothesis of Milton Friedman,<sup>21</sup> under which people's consumption decisions are based on their expected lifetime incomes, not on how much they earn in a given period. Section III also explores the significance for income averaging of modifying the assumption that distribution policy must rely

---

<sup>18</sup> Cites. Note that this assumes incomplete insurance markets.

<sup>19</sup> Cites.

<sup>20</sup> See, e.g., Mikhail Golosov, Aleh Tsyvinski, and Ivan Werning, *New Dynamic Public Finance: A User's Guide 1* (March 2006) (describing the static character of the standard Mirrlees approach).

<sup>21</sup> Milton Friedman, *A THEORY OF THE CONSUMPTION FUNCTION* (1957).

exclusively on static information about earnings, making use of an emergent branch of the economics literature, known as new dynamic public finance (NDPF), that is as yet little known to legal scholars. Section IV discusses the case underlying the new pro-consumption tax consensus, and shows how extensively it relies on the same assumptions as those that support income averaging, and thus is subject to similar objections. Section V offers a brief conclusion.

## II. THE OPTIMAL INCOME TAX FRAMEWORK

### A. Taxation and Inequality

Why would we have an income or consumption tax, let alone a tax with graduated marginal rates? A lump-sum tax, or one in which each taxpayer's liability is fixed without regard to any decisions that he or she makes,<sup>22</sup> would avoid distorting taxpayer behavior.<sup>23</sup> If this lump sum tax took the form of, say, a uniform head tax, or alternatively a reverse lottery to assign tax liabilities randomly, or even a tax based purely on people's eye color, it would not only be lump sum, but probably a lot simpler administratively than the current system, or indeed any that has seriously been proposed.

The reason for not having any such tax system is simple. Some people are better-off than others, and it is widely believed that those who are better-off should pay more tax. Thus, many would agree that Bill Gates should pay more than the average reader of this article, who in turn should pay more than a homeless person. This sometimes is called the criterion of ability to pay, perhaps reflecting assumptions about the effect of material wellbeing on the disutility of paying.<sup>24</sup>

In modern welfare economics, the notion of being better-off is commonly interpreted in terms of a budget line, reflecting the maximum combinations of available resources that an individual can acquire. Having a higher budget line translates to being better-off, all else equal, if we make the psychological assumption, standard in welfare economics, of non-satiation, i.e., that more of any good is always better than less.<sup>25</sup> Thus, if there are only two consumer goods, A and B, that are perfectly tradable for each other

---

<sup>22</sup> [Cite]

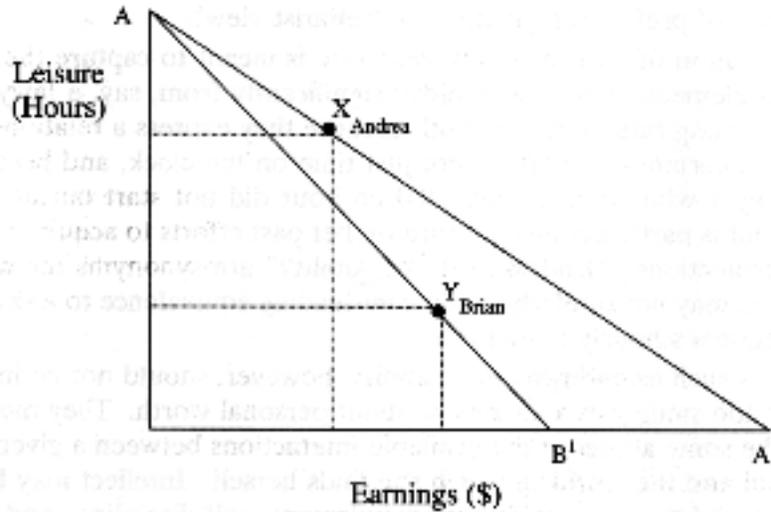
<sup>23</sup> Note substitution vs. income effects.

<sup>24</sup> [Cite]

<sup>25</sup> See Gruber, *supra*, at 25.

in complete markets, raising one's budget line means that one can get more of either or both without having to give up anything. Under non-satiation, this translates to being better, again holding all else equal.

Suppose that we assume two goods: (1) market consumption, comprising everything one could buy for cash, and (2) leisure, comprising not just free time but any use of one's time other than to earn as much money as possible. Everyone has twenty-four hours in a day, and without individuating information about people we might start out by assuming that everyone has an equal ability to make enjoyable use of time other than in trying to earn money. Clearly, however, people differ in wage rates, or ability to make money through economic production. This suggests a very simple framework for evaluating relative wellbeing, in which people's budget lines vary with their wage rates. The following is an illustration, in which Andrea has a higher wage rate than Brian, and therefore is assumed to be better-off even if, owing to different preferences, she decides to earn less.

FIGURE 1<sup>26</sup>

Even though Andrea earns less than Brian, we know under non-satiation that she must be better off, all else equal, because she could have earned the same amount while having more leisure, or alternatively had the same amount of leisure while earning more. She is at point X, rather than at a point that is equal to or better than point Y in both dimensions, simply because she prefers point X to any such point.

The conclusion commonly derived from this type of analysis is that, in principle, wage rate – not actual wages – is the key determinant of relative wellbeing that we have in mind when we reject lump sum taxation on the view that Bill Gates should pay more tax than we do while a homeless person should pay less. To be sure, if wages are all we can observe, as distinct from wage rate or labor supply, we will reach an erroneous conclusion in the above example, misclassifying Brian as better-off than Andrea and therefore as meriting less favorable treatment.<sup>27</sup> Yet this error should be unusual if we

<sup>26</sup> Figure 1 is taken from Daniel Shaviro, Endowment and Inequality, in Joseph Thorndike and Dennis Ventry (eds.), TAX JUSTICE RECONSIDERED: THE MORAL AND ETHICAL BASES OF TAXATION (2002).

<sup>27</sup> [And note the “beachcomber” problems with charging him less tax that I discuss in Endowment and Inequality.]

have no reason to think that wage rates generally are inversely correlated with preferences for market consumption. Even when we get correct relative rankings, however, it still remains to motivate responding to differences in wellbeing by imposing higher taxes on those, who by the admittedly imperfect rubric of actual earnings, appear to be better-off.

It should already be clear how the analysis reflects the assumption that earnings are the only observable variable that bears on relative wellbeing. The assumptions of complete markets and consistent rational choice are important as well. For example, we might be less sure that Andrea is better-off than Brian if gaps in the labor market prevented her from optimizing her work choice given her preferences.<sup>28</sup> Consistent rational choice tells us that people's labor supply decisions meaningfully reflect their preferences and opportunities, supporting viewing earnings information as generally meaningful even if, as in the case of Andrea and Brian, it occasionally supports mistaken conclusions.

#### B. Taxing Earnings in a Welfarist Framework

Merely observing evidence of unequal wellbeing does not immediately motivate responding to it with differential tax treatment. For this, one needs a normative framework. In welfare economics, the sole criterion for assessing a given policy is its effect on social welfare, which "is postulated to be an increasing function of individuals' wellbeing and to depend on no other factors."<sup>29</sup> In this framework, applauding Pareto improvements, where someone gains and no one loses, is easy but not immediately helpful in allocating tax burdens, an exercise that inevitably has winners and losers.

---

<sup>28</sup> See Shaviro, Endowment and Inequality, *supra* at \_\_\_\_.

<sup>29</sup> Louis Kaplow and Steven Shavell, FAIRNESS AND WELFARE 16 (2002).

Utilitarians assess these tradeoffs by requiring only that the net effect on welfare be positive, i.e., that gains exceed losses if one counts each individual's welfare equally.<sup>30</sup>

The most common alternatives to utilitarianism in welfare economics give extra weight to the welfare of worse-off individuals – at the limit, by giving infinitely greater weight to the welfare of the worst-off individual than to that of anyone else.<sup>31</sup>

Utilitarianism motivates redistribution from better-off to worse-off individuals through the assumption of declining marginal utility, which holds that unit  $n + 1$  of an item yields less utility than unit  $n$ .<sup>32</sup> If people have identical utility functions characterized by declining marginal utility, then transferring resources from better-off to worse-off individuals will increase social welfare, all else equal.<sup>33</sup> Variants of welfare economics that give extra weight to the welfare of worse-off individuals add an additional motivation for transferring resources to such individuals from those who are better-off.<sup>34</sup>

Within this framework, an earnings tax provides “an insurance mechanism to mitigate undesired risk from people's participation in the ability lottery.”<sup>35</sup> This requires departing from the complete markets assumption, since with complete insurance markets such coverage would already be available, and with consistent rational choice people would always choose to hold it if they wanted it. The gap in private markets that

---

<sup>30</sup> See *id.* at 5 n. 8.

<sup>31</sup> See Anthony Atkinson & Joseph Stiglitz, *Lectures On Public Economics* 339-340 (1980); Amartya Sen, *Equality of What?*, in 1 *The Tanner Lectures on Human Values* 206 (Sterling M. McMurrin ed., 1980).

<sup>32</sup> Shaviro, *supra*, at \_\_\_.

<sup>33</sup> In the case of Andrea and Brian, their different labor supply choices indicate that they do not have identical utility functions. Andrea evidently has less taste for market consumption, more taste for leisure, and/or greater work aversion. One might nonetheless treat Andrea and Brian, for purposes of the distribution decision, as if they had identical utility functions if this observed difference has no clearcut implications for the aggregate utility effects of the transfer. See *id.*

<sup>34</sup> Cite Shaviro, *Households and the Fiscal System*.

<sup>35</sup> Daniel Shaviro, *MAKING SENSE OF SOCIAL SECURITY REFORM* 52 (2000).

motivates government provision is most plausibly attributed to adverse selection, which arises when prospective purchasers of insurance have superior private information about their own prospects.<sup>36</sup> Adverse selection can prevent insurers from being able to offer a given product without losing money, by inducing disproportionate subscription by those who expect a positive payoff given their actual odds.<sup>37</sup> In the case of ability insurance, this would involve disproportionate subscription by those who knew they were likely to be low-income and thus to collect on the policies. The government can address adverse selection by requiring everyone to subscribe,<sup>38</sup> although the degree of socially desirable coverage remains limited by moral hazard, or unobserved behavioral responses to the incentive effects of the insurance coverage (here, by working less in response to a redistributive tax that is based on earnings).<sup>39</sup>

Following the early work of James Mirrlees,<sup>40</sup> economists have developed an extensive optimal income tax literature that relies on simulations, involving a hypothetical population and economy, to determine what the tax system should look like in furtherance of the ability insurance function. The main inputs to these simulations are the choice of social welfare function (utilitarian or otherwise), the distribution of earning ability in the population, the specification of people's utility functions, and the degree of labor supply responsiveness to the tax.<sup>41</sup> These typically are one-period models in which members of the hypothetical population do not face the question of whether to save any of their earnings for consumption in later periods. With everything being consumed

---

<sup>36</sup> Id. at 46.

<sup>37</sup> See Gruber, *supra*, at 311-314, note death spiral.

<sup>38</sup> Id. at 314.

<sup>39</sup> Shaviro, *MAKING SENSE OF SOCIAL SECURITY REFORM*, *supra*, at 55.

<sup>40</sup> James A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38(2) *Rev. Econ. Studies* 175 (1971)..

<sup>41</sup> See, e.g., Gruber, *supra*, at 570-572..

currently, earnings, income, and consumption are all the same. This eliminates any consideration either of the choice between income and consumption taxation or of income averaging, which would require multiple periods. The simulations typically focus on what an optimal tax rate structure would look like, thus helping to inform our judgment about tax rate structures in the real world if we are sufficiently sanguine about the models' representativeness as a simplified version of the real world. The usual conclusions reached are that everyone should receive a minimum grant (regardless of income) and that marginal rates, as income rises, should be surprisingly non-graduated.<sup>42</sup>

Another important result in the broader welfare economics literature comes from Anthony Atkinson and Joseph Stiglitz, who showed that, if distinctive taxes on specific commodities can be imposed to advance redistributive aims, in general all commodities should be taxed at the same rate.<sup>43</sup> In a one-period model where earnings, income, and consumption are the same, simply taxing earnings is equivalent to imposing the same tax on all commodities (i.e., everything on which the earnings might be spent). The underlying intuition is that any extra tax on a given commodity discourages labor supply, since it reduces the amount of consumption one can purchase by working, and that differential commodity taxation therefore simply adds a further distortion, concerning people's consumption choices. This new distortion is layered on top of the other, increasing total distortion, rather than in any way mitigating it.<sup>44</sup> As we will see, this analysis has implications both for ideal tax base debate and for income averaging.

---

<sup>42</sup> Cites, explain reason for this result & the result of a zero rate at the very top.

<sup>43</sup> A.B. Atkinson and J.E. Stiglitz, *The Design of Tax Structure: Direct Versus Indirect Taxation*, 6 *J. Pub. Econ.* 55 (1976). This result requires that all commodities be weakly separable from leisure, i.e., that none be leisure substitutes or leisure complements, the consumption of which was differentially affected by how much leisure one had chosen.

<sup>44</sup> Cite Bankman & Weisbach.

### C. Extensions of the Mirrlees Framework

If ability is the main attribute of distributional interest from the standpoint of welfare economics, and earnings are of interest merely as evidence of ability, there is no reason to assume that earnings are the only evidence worth observing. For example, one could in principle also try to make use of information about gender, ethnic origin, or observable physical characteristics such as height if determined to be statistically relevant.<sup>45</sup> As yet, few practical suggestions have been made in this area apart from treating disability as evidence of low earning capacity.<sup>46</sup>

One should keep in mind, however, that even ability matters only as evidence bearing on total and marginal utility. An individual in a permanent coma, while having no earnings ability, would presumably lack appeal to a utilitarian as a prospective recipient of large transfers.<sup>47</sup> Ability as a consumer – that is, the capacity to derive extra utility from extra resources – matters distinctly from earning ability. Differences in ability as a consumer are usually ignored, however, reflecting the difficulty of observing them along with the analytical advantages of using a simpler framework.<sup>48</sup>

The static one-period approach of classic optimal income tax models likewise serves purely to simplify the analysis, concededly at the cost of accuracy and descriptive richness. Over the last twenty years, the NDPF literature has begun incorporating time, risk, and gradually unfolding information into the Mirrlees framework.<sup>49</sup> While little known as yet to legal scholars, this literature has important implications for income averaging and choice of tax base, as we will see in the next two sections.

---

<sup>45</sup> [Note studies saying height is correlated with income or with IQ]

<sup>46</sup> Note that fiscal system does this, and cite Sanchirico on “clumsiness” in relation to tort judgments.

<sup>47</sup> [Cite Kaplow?]

<sup>48</sup> Note the “utility monster” issue if we take into account differences in ability as a consumer.

<sup>49</sup> [Cites]

### III. THE CASE FOR (AND AGAINST) INCOME AVERAGING

#### A. Annual Versus Lifetime Systems

The mainly annual character of the federal income tax – shared by comparable tax systems around the world although not by Social Security and similar retirement systems – has two distinct components, apart from the frequency of filing. The first pertains to information, and the second to the timing of cash flows between the taxpayer and the government.

As to information, tax liability largely depends on measuring the taxpayer's position during the current year. For the most part, one's income for the year determines one's tax liability without regard to one's income in other years, leaving aside the carryover of various tax attributes such as net operating losses.<sup>50</sup> As to the timing of cash flows, taxpayers generally must settle up each year with the Internal Revenue Service. They cannot, at a market interest rate, accelerate or defer the tax payments that are due. Early payment is not rewarded with a time value discount to keep its present value constant, and late payment (beyond permissible extensions) is penalized, rather than simply leading to the imposition of a market interest charge. Poverty programs likewise rely mainly on current year (or even current month) information, with exceptions such as the five-year time limit on receiving Temporary Aid to Needy Families (TANF) benefits.

The restriction to current-year information inevitably has important effects in the presence of non-flat marginal rates (including not only rising rates, but those that decline with income due to phase-outs, such as those of income tax credits or welfare benefits). If non-flat rates apply to annual income, then the sequence of one's earnings and other taxable income, and in particular whether they fluctuate or are relatively constant over

---

<sup>50</sup> However, attributes such as asset basis make possible the use of information from past years.

time, affects one's lifetime tax burden. Earning fluctuations can significantly increase one's lifetime tax burden if marginal rates rise with income.

By contrast, the requirement of annual cash flow settlement matters only contingently. In the presence of (1) complete capital markets, permitting people to borrow and lend across time however they like, and (2) rational consumers who make consistent intertemporal choices given their preferences, it would make no difference whatsoever. Having to pay federal income tax at a given time, rather than being allowed to accelerate it or defer it at a market interest rate, would have no effect on people's work or consumption activity if they could (and did) borrow and lend at will to arrange their overall cash flows as they liked. However, where markets are incomplete or people fail to exercise consistent rational choice across time, the sequence of cash flows between oneself and the government can be important.

Historically, the first of these two features of an annual tax system has been a lot more controversial than the second. In particular, it has led to calls for income averaging, under which marginal rates would apply to one's average annual income over a period longer than a year – perhaps even one's entire life. Annual cash settlement, by contrast, has prompted little dissatisfaction. While caring about it may be logically reconcilable with favoring the use of a long-term income picture,<sup>51</sup> the two are uncomfortable bedfellows. For annual cash settlement to matter, the current sequence of cash flows – deemed inconsequential in the long-term view – must matter.

---

<sup>51</sup> For example, one might favor annual cash settlement on the ground that the government is ill-equipped to handle the default problems that might be raised by letting people borrow their current tax liabilities at a market interest rate.

This ambivalence is prominently on display in the best-known and most comprehensive income averaging proposal to date, made by economist William Vickrey in 1939.<sup>52</sup> The normative criteria on which Vickrey relied included the following:

- (1) The discounted value of the series of tax payments made by any taxpayer should be independent of the way in which his income is allocated to the various income years. ...
- (4) Any given tax payment should not be too large in relation to the income of the period immediately proceeding (Vickrey 1994 ed. 107).

These two criteria can be (and in Vickrey's scheme are) pursued simultaneously, since the former relates to lifetime tax burden and the latter to the timing of tax payments. Yet Vickrey's Criterion (4) is a poor intellectual fit with his Criterion (1). If only the discounted present value of cash flows matters, why worry about the relationship between a given tax payment and income of the immediately preceding period? And if we do need to worry about that relationship, doesn't this suggest that not all present-value-equivalent cash flows are the same?

The reasons for this ambivalence lie in the explanatory power and limits of economic theory. Vickrey's Criterion (1) would be completely persuasive, and his Criterion (4) irrelevant, if one fully accepted the permanent income hypothesis of Milton Friedman, under which people's consumption decisions are based on their lifetime incomes, not on how much they earn in a given period.<sup>53</sup> However, given incomplete capital markets and time-inconsistent preferences, the permanent income hypothesis does not fully hold. As I discuss next, its descriptive validity is context-dependent in ways

---

<sup>52</sup> I discuss the Vickrey plan *infra* at \_\_\_.

<sup>53</sup> Friedman, *supra*..

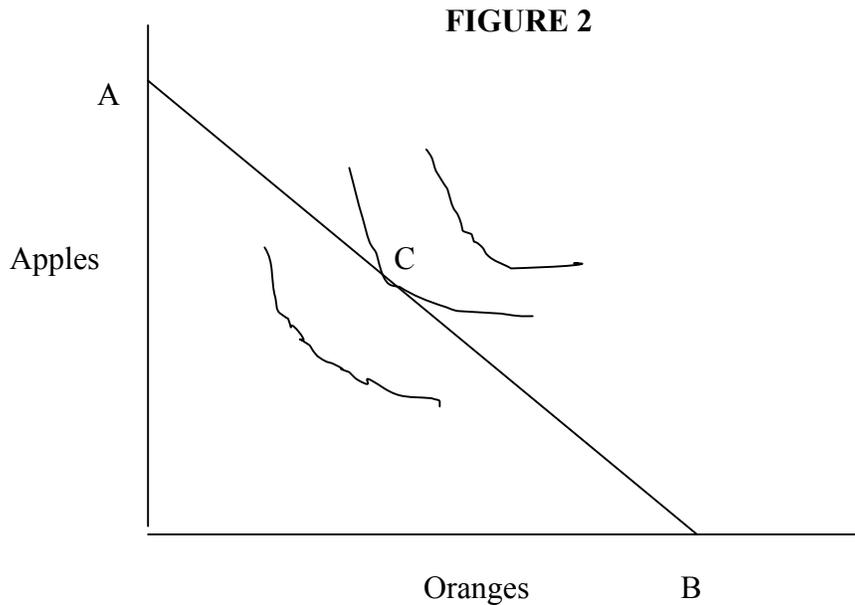
that the government cannot easily observe, and that general tax rules cannot easily capture.

B. The Permanent Income Hypothesis

The Friedman permanent income model is mathematically sophisticated. So is the closely related life-cycle model of consumption smoothing that was pioneered by Franco Modigliani and Richard Brumberg.<sup>54</sup> Yet the core idea behind both models is simple and intuitive, reflecting familiar economic reasoning. To illustrate it in a different setting, suppose we again posit a world with only two consumer goods, here apples and oranges, which are freely tradable for each other at a fixed ratio and at zero transaction cost. Once again, each worker has a wage rate and therefore a budget line, reflecting the largest combinations of the two consumer goods that she can earn. Here, however, I address how one chooses a particular combination of the two. One's choice depends on one's preferences, which are assumed to reflect declining marginal utility for each good, and which can be represented through indifference curves showing combinations of the two goods that one rates as equal, as shown in Figure 2.

---

<sup>54</sup> See Franco Modigliani and Richard Brumberg, *Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data*, in K.K. Kurihara, ed., *POST-KEYNESIAN ECONOMICS* (1954). See also Angus Deaton, *UNDERSTANDING CONSUMPTION* (1992) at 214, describing the permanent income and life cycle hypotheses as “well-defined special cases of the general theory of intertemporal choice.” For convenience given that the differences between the two models are unimportant for my purposes, I emphasize Friedman's permanent income hypothesis throughout.



The worker in Figure 2, who can obtain any of the apple-orange combinations on line AB, picks the combination at point C, which lies on the highest indifference curve that is tangent to AB.

Now suppose that, despite the free exchangeability of apples and oranges, some of the firms that might employ this worker pay purely in apples, while others pay purely in oranges. Under standard economic reasoning, the worker would be expected to end up at C no matter which firm ends up employing her. Given the goods' free exchangeability, and adding as well the assumption of consistent rational choice given one's preferences, she will trade her way to the favored spot even if she starts at A or B. Under these assumptions, we also would expect her choice of employer to be unaffected by whether a given firm paid in apples or in oranges.

A further possible implication requires more assumptions. Suppose that, if two workers at the same budget line chose the same commodity mix, we would assume that they had the same total utility and marginal utility for an extra unit of consumption. We might base this assumption on the view that, so far as we can tell, they have identical utility functions. By contrast, if we observe two workers making different choices at the same budget line, we learn that their utility functions must be different. Still, if this observation conveys no information to us about the two workers' total or marginal utility, we may reasonably continue to treat the two workers as identical.

This little model has wide-ranging applicability. Make the goods “market consumption” and “leisure,” with tax being imposed only on the former, and you have the basic set-up for optimal income tax analysis (as discussed in section II), with its implication that the equal taxes imposed on people who are relevantly equal should in principle reflect ability (i.e., the wage rate), rather than labor supply choice.<sup>55</sup> Emphasize the assumption that the worker will head to point C no matter where she starts out, and essentially you have the Coase theorem.<sup>56</sup> Or, of greatest interest here, make the two goods “earlier consumption” and “later consumption,” and you get the permanent income hypothesis.

The permanent income hypothesis holds that people's current consumption choices depend on their anticipated lifetime income, rather than their income for any given period. Thus, the sequence of one's earnings is entirely independent of the sequence of one's consumption, keeping in mind that labor effort (as distinct from earnings) in a given period may affect how much one wants to consume in that period.

---

<sup>55</sup> See, e.g., Daniel Shaviro, Endowment and Inequality, in Joseph Thorndike and Dennis Ventry (eds.), *TAX JUSTICE RECONSIDERED: THE MORAL AND ETHICAL BASES OF TAXATION* (2002).

<sup>56</sup> See Ronald Coase, *THE FIRM, THE MARKET, AND THE LAW* (1988).

While not logically necessary to the model thus described, proponents note that in practice it is likely to imply smoothing out one's annual consumption stream relative to one's annual earning stream.<sup>57</sup> In particular, to consume after retirement, people must save for it. Likewise, good years and bad years from an earnings standpoint need not have any correlation with years when one prefers high consumption as opposed to low consumption, suggesting a possible need for consumption smoothing even during one's working years.

If consumption in any one period has declining marginal utility as the amount of consumption in that period increases, people will have some tendency to prefer complete income smoothing as between periods. There is no reason to predict complete smoothing, however, with realistic utility functions. Various plausible preferences, such as for rare but long vacations, or to travel more while one is still young, can lead to one's choosing uneven consumption between periods, or particular patterns such as smoothly rising or falling levels of consumption.

### C. The Case for Income Averaging Under the Permanent Income Hypothesis

Under the permanent income hypothesis, the case for income averaging is compelling on both distributional and efficiency grounds, ignoring for now the possibility that earnings or consumption sequences, even if rationally chosen in complete markets, might offer pertinent information about ability.

#### 1. Distribution

In the permanent income model, people's consumption over time depends purely on their budget lines, lifespans, and intertemporal consumption preferences. Thus, two individuals with different earnings sequences but identical lifespans, budget lines, and

---

<sup>57</sup> See, e.g., Lee Anne Fennell and Kirk Stark, *Taxation Over Time* (forthcoming in *Tax Law Review*).

consumption choices would have identical total and marginal utility, so far as we know, if the differences in their earnings sequences did not convey information to the contrary.

Under a utilitarian social welfare function, the two individuals' presumed identical marginal utility would support taxing them the same. If one shifted to a social welfare function that gave extra weight to the welfare of worse-off individuals, one still would want to tax them the same, given their presumed identical total utility.

An important detail here concerns lifespan differences. The longer one lives, the greater one's lifetime consumption needs. Thus, while living longer may increase one's total utility, it also increases the marginal utility of an extra dollar. In the absence of complete markets and consistent rational choice, this would suggest that, as between two individuals with the same lifetime income, the one who lives longer should pay less in lifetime net taxes. No government response would be necessary, however, if people fully responded to life expectancy risk by annuitizing – that is, by using their wealth, minus any amounts they wanted to leave in bequests, to purchase life annuities that would support them for as long as they lived.

## 2. Efficiency

Whether the tax system is based on income, earnings, or consumption, the use of a lifetime measure improves efficiency, assuming non-flat marginal rates, because it causes the same marginal rate to apply to the taxpayer's choices in all years. Suppose initially that tax liability depends on income or earnings. Using an annual rather than a lifetime measure improves efficiency in two respects. First, with a shorter than lifetime measure, people have an incentive to shift their earnings to years where they face a lower rather than a higher marginal rate. A lifetime measure avoids this. Second, a lifetime

measure permits tax rate smoothing. One can raise the same revenue from a given individual as under an annual system by blending what would otherwise be some years' high rates and other years' low rates into a single intermediate set of rates. This would be expected to reduce economic distortion even if earnings cannot be shifted between years, because the increase in distortion as rates increase is more than linear.<sup>58</sup> Thus the reduced distortion from lowering the rates in some years should exceed the increased distortion from raising them in other years.<sup>59</sup>

Now suppose instead that tax liability depends on consumption. The use of a lifetime measure results in what Atkinson and Stiglitz would classify as a uniform commodity tax.<sup>60</sup> One faces the same marginal rate on an extra unit of consumption no matter the year in which it occurs. If the applicable marginal rate depends instead on the taxpayer's marginal rate bracket given her other consumption in a particular year, we get a violation of the Atkinson-Stiglitz analysis. In effect, commodity taxes are higher at the margin in some years than in others, distorting decisions regarding when to consume without any mitigation of the basic underlying labor supply distortion.<sup>61</sup>

#### D. Implementing Income Averaging

William Vickrey's "cumulative averaging" scheme remains the fullest proposed implementation of the idea that lifetime tax liability should depend on permanent

---

<sup>58</sup> Jonathan Gruber, PUBLIC FINANCE AND PUBLIC POLICY 552-553 (2005).

<sup>59</sup> Weisbach, The Optimal Accounting Period for Taxes (unpublished manuscript) emphasizes two considerations potentially raised by income averaging that I ignore here. The first is that longer accounting periods increase economic distortion if there is no time value adjustment for when in the period one received a given dollar. In principal, however, no matter how long or short the period one uses informationally, one can make present value adjustments within the period. Second, Weisbach notes the administrative costs of requiring more frequent filing, clearly an important consideration although distinct from what information is used each time and from what payment schedule is imposed.

<sup>60</sup> Cite AS 1976, and note that they allow rate graduation for direct taxation as a whole.

<sup>61</sup> Cite Bankman & Weisbach..

income.<sup>62</sup> Vickrey's basic idea was to have the taxpayer, each year, determine her average annual income through that year, and then adjust the amount she had paid to date to equal what she would have paid to date had she earned the average amount each year.

A simplified illustration of the Vickrey proposal, assuming for simplicity an interest rate of zero, is as follows. Suppose the tax rates are 20 percent for income up to \$100,000 and 50 percent above that, and that only two years are at issue. In addition, suppose that three taxpayers (A,B, and C) each earned \$200,000 total, but in different sequences: A had an even split between the two years, B earned everything in Year 1, and C earned everything in Year 2. The Vickrey system would apply as follows:

**TABLE 1**

	YEAR 1		YEAR 2		TOTAL	
	Earn	Tax	Earn	Tax (Refund)	Earn	Tax
A	\$100,000	\$20,000	\$100,000	\$20,000	\$200,000	\$40,000
B	\$200,000	\$70,000	0	(\$30,000)	\$200,000	\$40,000
C	0	0	\$200,000	\$40,000	\$200,000	\$40,000

(Tax rate is 20% on income up to \$100,000, 50% above that;  
B and C pay in Year 2 the amounts needed to equalize them with A)

Vickrey did not definitely commit himself to a particular averaging period, but noted the logic of extending it all the way from adulthood (reflecting that few children have significant earnings) until death.<sup>63</sup> This presumably would cause retirees to get annual cash refunds for as long as they lived, since, in the absence of significant current year income, the reduction in average earnings would be treated as reducing taxes due for all prior years. Vickrey's system might therefore effectively provide retirees with life annuities, albeit declining ones given that each year the arithmetic effect of adding one

<sup>62</sup> See Fennell and Stark, *supra*, at \_\_\_\_.

<sup>63</sup> William Vickrey, *AGENDA FOR PROGRESSIVE TAXATION* 186 (1947).

more year to the denominator in the average earnings computation would decline slightly.<sup>64</sup> This annuity feature may be irrelevant, however, if one assumes complete markets and consistent rational choice, since under those assumptions people would arrange for the exact level of annuitization that they wanted, whether income averaging contributed to it or not.

Vickrey's system goes well beyond anything one could realistically expect Congress to enact. However, from 1964 through 1986, the U.S. tax system had a narrower income averaging rule, under which people with rising incomes could in some circumstances apply the lower marginal rates that they had faced in preceding years to some of the growth component. In terms of Table 1, relief was offered to individuals resembling C, but not to those resembling B. Income averaging was repealed in 1986, not on the view that an annual system is best, but rather on the grounds (in considerable tension with each other) that (1) rate reduction made it unnecessary, and (2) the revenue gain from repeal was needed to help pay for the reform.<sup>65</sup> So the main remaining form of income averaging in the U.S. system is that resulting from the allowance of net operating losses.

A second income averaging episode worth noting is that of Wisconsin, which between 1929 and 1935 based income tax liability on the taxpayer's average annual income for the prior three taxable years.<sup>66</sup> The rule applied automatically rather than being elective, and it benefited taxpayers with falling as well as rising income.

---

<sup>64</sup> [Explain & illustrate. Note that, since Vickrey used an income measure, one might conceivably match average annual income in a retirement year.]

<sup>65</sup> See Joint Committee on Taxation, General Explanation of the Tax Reform Act of 1986.

<sup>66</sup> See Lily L. Batchelder, Taxing the Poor: Income Averaging Reconsidered, 40 Harv. J. Leg. 395, 415 n. 686 (2003); Vickrey, *supra*, at 183-184.

One important difference between Vickrey-style cumulative averaging and the former U.S. and Wisconsin systems concerns the direction of their effects on current year tax payments when annual income fluctuates. Under Vickrey averaging, rising income tends to increase the current tax payment due, while falling income tends to reduce it. This reflects that one is doing more than just determining current year tax liability based on average income over a longer period. One also is in effect adjusting the tax payments due for prior years, by reconstructing what total taxes one would have paid for all years with level annual income. Thus, rising income effectively increases the taxes due for past years, potentially leading to an extra net levy payable in the current year, and falling income works in the opposite direction.

By contrast, the U.S. and Wisconsin systems treated only current year liability as open for adjustment based on past information. The consequence was that rising income tended to reduce current year tax liability as a percentage of current income, while falling income (if the system addressed this scenario, like the Wisconsin system) increases the percentage paid in the current year. Vickrey criticized such a relationship between income fluctuation and current year tax liability, arguing that, if one “assess[es] a heavy tax based on previous high incomes in years when the income has sharply decreased ... collection is difficult and hardship to the taxpayer results.”<sup>67</sup> While intuitively plausible for reasons that I discuss below,<sup>68</sup> this claim is in tension with the view that lifetime income, without regard to the exact sequence of cash flows, is all that matters.

E. Problems With the Case for Income Averaging

---

<sup>67</sup> Vickrey, *supra*, at 169.

<sup>68</sup> See text accompanying notes \_\_\_, *infra*.

The twin assumptions of complete markets and consistent rational behavior may have been easy to accept in the context of a stylized hypothetical concerning a world with nothing but apples and oranges. The assumptions become a lot more heroic, not to mention implausible and demonstrably untrue, when we shift to real world questions of the timing of consumption across one's lifespan.. In addition, with or without the twin assumptions, the case for income averaging may be weakened by the significance of additional information, whether revealed by the taxpayer's earning or consumption sequences or by our general knowledge about people's lifecycle patterns.

### 1. Incomplete Markets

In the hypothetical with the apples and oranges, all that the "complete markets" assumption required was full tradability between these two goods. Once we turn to lifetime income, however, the complete markets that one needs are capital markets. People must be able to (1) invest current earnings for future consumption, (2) borrow against future earnings for current consumption, and (3) adjust fully and immediately to changes in current earnings whether anticipated or not. Only the first of these three requirements is relatively unproblematic. Departures from the latter two include the following:

a) "Great expectations" – Borrowing against high expected future earnings is not always feasible, even absent risk concerning a given individual's capacity to realize her "great expectations" in the future. For example, students at leading professional schools who are virtually guaranteed future high-wage employment opportunities may face significant limits on their capacity to borrow against that future capacity to finance current consumption, even if they can borrow extensively to fund their educations.

These difficulties reflect the classic insurance problems of moral hazard and adverse selection. As to moral hazard, once I have borrowed against expected future earnings but will be judgment-proof unless I actually realize those earnings, the liability functions like a tax on the earnings, diminishing my incentive to realize them. As to adverse selection, I may know more than prospective lenders about the future earnings I actually should expect. This may reflect inside information either about my ability level or about my future plans.<sup>69</sup>

To illustrate the significance of the great expectations problem for income averaging, suppose that Caleb and Diana face identical circumstances, including having the same lifetime income, except that Caleb has level earnings once he achieves adulthood, while Diana has back-loaded earnings that she cannot access during the “great expectations” stage. Suppose further that Caleb and Diana both prefer perfect consumption smoothing during their adult years. The fact that Caleb can accomplish this while Diana cannot, given available capital markets, rebuts the presumption, derived from the permanent income hypothesis, that they are equally well-off. Caleb, as compared to Diana, is (a) better-off on a lifetime basis, given his ability to achieve the preferred consumption pattern, (b) better-off during the early adult years when Diana, unlike him, is effectively poor, and (c) worse-off once Diana’s high earning period has begun. At this later stage, while Diana still is worse-off on a lifetime basis given the earlier deprivations that she would have avoided if possible, she now has extra cash available and, for the remaining period, is effectively richer than Caleb.

---

<sup>69</sup> Education loans may be less risky than loans for general consumption purposes, notwithstanding the moral hazard and adverse selection problems, because I may be less likely to want access to funds that are being spent on the education if I do not actually plan to use it towards the goal of realizing high earnings.

Under these circumstances, the distributional equity argument for income averaging fails to hold. In a sense, lifetime comparisons between Caleb and Diana are not even meaningful, since the relative value of a dollar to either of them depends on when it is realized (a point that the lifetime income measure ignores). Rather than ensuring that they pay the same lifetime taxes, as would a perfect income averaging system, we should want to transfer resources from Caleb to Diana before her great expectations begin to be realized, and from Diana to Caleb afterwards. This suggests treating Diana's two periods as separate rather than as subject to being income-averaged with each other, so that she will (rightly) appear to be a low-income taxpayer in the first period and a high-income taxpayer in the second one.

The efficiency arguments for income averaging likewise lose ground in the great expectations scenario. Level lifetime rates continue to be desirable from the standpoint of eliminating incentives to shift earnings between periods. Moreover, there is still some argument for rate smoothing insofar as (all else equal) lowering high rates tends to reduce distortion more than raising low rates increases it.<sup>70</sup> On the other hand, Diana is, in effect, two different individuals across time even if she has constant preferences. She is a low-earner early on and a high-earner later on. This might lead to differences in her labor supply elasticity as between the two periods, since neither the rate of trade that she faces between consumption and leisure nor her budget line are the same. Such differences might cause smooth rates to be less efficient than having a higher marginal rate in the period (whichever it is) in which her labor supply is less responsive to changes in her after-tax return. As noted in a recent contribution to the new dynamic public finance

---

<sup>70</sup> Cite Rosen or Gruber.

(NDPF) literature, when one's skill level changes, "the tradeoff between insurance and incentives then shifts and taxes should adjust accordingly."<sup>71</sup> Tax rate smoothing ignores this.

b) Risk – A second, less obvious missing capital market relates to risk and consequent changes in information when a risk is resolved. If I expect low career earnings and then suddenly learn that they will be high or vice versa, I cannot (absent a time machine) suitably adjust the amounts I consumed in earlier periods to reflect what I now know. Accordingly, even if all other conditions for the model's applicability fully hold, I will fail ex post to achieve the optimal sequencing of my lifetime consumption.

The private market solution to this problem would be for me to fully insure against risks that my earning ability will change. However, moral hazard and adverse selection prevent full insurance, or anything close to it, from being available. A lesser mitigation device, precautionary saving against the risk of a decline in actual or expected earnings, increases expected smoothing but may actually end up reducing it ex post.

The classic permanent income and consumption smoothing models exclude uncertainty,<sup>72</sup> which Milton Friedman conceded "blurs the sharp lines of the analysis and suggests additional factors that may produce departures from the shape of the consumption function [otherwise] specified."<sup>73</sup> While merely an additional technical challenge from the standpoint of properly specifying models to describe rational lifecycle behavior,<sup>74</sup> uncertainty more definitely compromises reliance on the permanent income

---

<sup>71</sup> Golosov, Tsyvinski, and Werning, *supra*, at 17.

<sup>72</sup> See Friedman, *Theory of the Consumption Function*, *supra*, at 7-14; Modigliani and Blumberg, *supra*, at 392.

<sup>73</sup> Friedman, *supra*, at 15. He had particularly in mind precautionary saving. See *id.* at 16.

<sup>74</sup> See Friedman, *supra*, at 16 (suggesting that uncertainty would encourage precautionary saving); Modigliani and Blumberg, *supra* at Modigliani and Blumberg, *supra* at 392 (arguing that "a satisfactory

hypothesis to support viewing all individuals who ex post have the same lifetime incomes as having faced identical circumstances.

Uncertainty undermines both the distributional and the efficiency cases for income averaging. Suppose that two periods in a taxpayer's life differ, so far as earnings are concerned, in part predictably (e.g., one's income rises with skill or seniority) and in part unpredictably due to the resolution, at the dividing line between two periods, of a discrete risk. The predictable difference would lead a rational individual to engage in consumption smoothing between the two periods, suggesting that income averaging would be appropriate unless smoothing was impeded by the "great expectations" problem. However, inability to retrofit the amount of earlier-period consumption to resolution of the risk would suggest treating the two periods as separate, and thus not permitting income averaging so far as the risky outcome is concerned.

Risk disrupts basing distributional judgments purely on lifetime income even if two individuals turn out ex post, not only to have had the same total lifetime income, but even to have earned exactly the same amount each year. Suppose Edith and Frank end up earning the same amount each year, but that Edith is subject to greater earnings risk throughout and thus rationally chooses greater precautionary saving. She thus ends up back-loading her lifetime consumption relative to Frank's (all else equal) and relative to what would have been optimal for her had her actual ultimate earnings stream been certain. Frank therefore experiences greater total utility, all else equal, possibly motivating redistribution from him to Edith even though, judged through the lens of income averaging, they end up being the same.

---

theory can be developed without seriously coming to grips with this rather formidable problem"), and at 428-429 (noting some ways in which uncertainty may affect saving ).

A further risk-related distributional issue, arising from incomplete markets, concerns difficulties in promptly adjusting consumption levels in response to changes in actual and expected earnings. Suppose a high-earner who owns an expensive house suddenly and unexpectedly loses her job, indicating that she must start to live somewhere more modest. Adjusting her housing consumption may take time (e.g., since she needs to find a buyer or renter), all the more disruptively if she is carrying a mortgage with high monthly payments that she cannot finance out of other savings. Vickrey presumably had such scenarios in mind when he criticized Wisconsin-style income averaging for responding inappropriately to declining income. He appears not, however, to have appreciated the broader tension between such scenarios and viewing lifetime income as the canonically correct basis for determining lifetime tax burdens.

Turning to efficiency, the change in skill level likewise undermines the case for income averaging. Once again, in light of the change “the tradeoff between insurance and incentives ... shifts and taxes should adjust accordingly.”<sup>75</sup>

c. Inability to annuitize – Recent research suggests that life annuities generally are not available at actuarially fair prices, presumably due to adverse selection.<sup>76</sup> That is, people with private information suggesting that they will live longer than indicated by the actuarial tables sign up disproportionately, making fair returns

---

<sup>75</sup> Golosov, Tsyvinski, and Werning, *supra*, at 17.

<sup>76</sup> See Amy Finkelstein and James Poterba, Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market, 112 *J. Pol. Econ.* 183 (2004); James Poterba, Steven Venti, and David Wise, Personal Retirement Savings Programs and Asset Accumulation: Reconciling the Evidence, in David Wise (ed.), *FRONTIERS IN THE ECONOMICS OF AGING* (1998). If life annuity markets are badly hampered by adverse selection, it is unclear why markets for life insurance are so robust, given that the two financial instruments reflect the same bet (whether the insured will die early or late), differing only in the side of the bet that the insured takes.

based on those tables a losing proposition for the insurers. This gap in financial markets arguably suggests that the government should provide mandatory life annuities.<sup>77</sup>

Under this view, the life annuity aspect that income averaging features when the averaging period continues through death may be a step in the right direction, rather than being irrelevant as suggested above. Given, however, that other instruments for mandatory annuitization exist, such as mandatory retirement programs along the lines of Social Security and Medicare, the significance of income averaging's contribution here remains uncertain. One presumably needs to coordinate it with the other programs, but it does not expand the government's opportunity set, nor does its design have anything directly to do with the question of how much mandatory annuitization is desirable.

## 2. Departures from Consistent Rational Choice

A second empirical failing of the permanent income hypothesis is its assuming consistent rational choice. Behavioral economics research has revealed a number of real world departures from this model of behavior.<sup>78</sup> These departures suggest that two people with the same lifetime incomes but in different sequences, and with the same preferences may make different consumption choices, implying that they will face different circumstances, even with complete markets.

a. Hyperbolic discounting and other causes of myopia – There is considerable evidence that people often fail to save adequately for retirement, reflecting myopia rather than a consistent preference for concentrating consumption in their working years.<sup>79</sup> Psychological explanations differ, but include hyperbolic discounting,

---

<sup>77</sup> Gruber, *supra*, at 343.

<sup>78</sup> See, e.g., Cass R. Sunstein (ed.), *BEHAVIORAL LAW AND ECONOMICS* (2002); Richard H. Thaler, *QUASI-RATIONAL ECONOMICS* (1991).

<sup>79</sup> Cites

or applying a much higher discount rate between the current time and any future time than between future times.<sup>80</sup> People who are subject to hyperbolic discounting cannot hold consistent preferences. For example, at Time 1 a hyperbolic discounter will want to apply a normal discount rate in dividing consumption between Times 2 and 3. Once Time 2 arrives, however, she will be much more inclined to concentrate her consumption in Time 2.

To the extent that people act under the sway of hyperbolic discounting, current consumption will depend on currently available resources rather than on any consistently held long-term plan. To take an extreme case, suppose people always consumed all current earnings, neither saving for the future nor borrowing against the present value of expected future earnings. Then, for any year, both total utility and the marginal utility of a dollar would depend purely on current earnings (ignoring psychological carryover effects from remembering past years or anticipating future ones), and income averaging would give current distributional weight to information from other periods that was in fact irrelevant. Even under a less extreme view, to the extent that people lean towards current consumption of currently available resources, not because they are optimizing across the lifespan but because they are behaviorally subject to “presentist” bias, current year information gains distributional relevance relative to information from other periods, thereby weakening the case for income averaging.

This point has particular implications for the design of need-based transfers such as TANF. Proponents of income averaging do not generally take the intellectually consistent step of urging that eligibility for need-based transfers depend purely on lifetime income, which would suggest denying the benefits to currently destitute

---

<sup>80</sup> See, e.g., David Laibson, Golden Eggs and Hyperbolic Discounting, 112 Q.J. Econ. 443, 445-446 (1997).

individuals who had high earnings in the past. This departure from consistently favoring a lifetime standard can be defended on the ground that being destitute despite past earnings is prima facie evidence of poor lifetime consumption choice, rather than of having rationally preferred to concentrate one's consumption in earlier periods.

b. Mental accounts – Even without thoroughgoing myopia, consumption behavior may diverge from that predicted by the permanent income hypothesis due to people's use of "mental accounts" to classify dollars differently depending on their source. In effect, these are rules of thumb that people use to economize on decision costs. Thus, amounts coded as "current income" apparently are more likely to be spent than those coded as "current assets," which in turn are more likely to be spent than those coded as "future income."<sup>81</sup> Empirical research strongly supports the view that these predilections cause significant behavioral departures from the predictions of the permanent income hypothesis, in particular by causing current consumption and current income to be "much more closely linked" than one would expect if it fully applied.<sup>82</sup> Once again, the implication is that the time sequence of earnings matters independently of their discounted lifetime value, and thus that current year information should have greater influence on current year distribution policy than information from other periods.

### 3. Additional Information

The problems with income averaging that I have discussed thus far seem paradoxically to suggest that it uses too much information, by going beyond current year earnings to look at other years' earnings as well. What makes this seem paradoxical is

---

<sup>81</sup> See Hersh M. Shefrin and Richard H. Thaler, The Behavioral Life-Cycle Hypothesis, in Thaler (ed.), QUASI RATIONAL ECONOMICS 96-98 (1991).

<sup>82</sup> Christopher C. Carroll & Lawrence H. Summers, Consumption Growth Parallels Income Growth: Some New Evidence, in B. Douglas Bernheim and John Shoven (eds.), NATIONAL SAVING AND ECONOMIC PERFORMANCE (1991).

that one would think that more information is always better than less, particularly when the underlying distribution problem is framed in information terms as one of trying to gauge individuals' true ability levels.

The explanation for the apparent paradox is as follows. Reliance on the permanent income hypothesis involves treating information from other periods not just as currently relevant, but as equal in relevance to current period information. It flattens the lifespan, treating it as a single period and ignoring distinctions between stages, along with details concerning the exact sequences of earnings and consumption. When the internal pattern is important, a purely annual approach, while clearly suboptimal unless the other periods are totally irrelevant, can potentially be better than acting as if all information from all periods was equally relevant at all times. Better to ignore the other information, perhaps, than to over-weight it.

One can also, however, take the route of improving the lifetime measure by adding information, in lieu of subtracting it. Rather than treat the lifetime as a single period, one can instead attempt proper relative weighting and use of the information from different periods. Here the NDPF literature becomes especially pertinent, as it expressly focuses on the significance of information that unfolds over time.<sup>83</sup> Examples of using information beyond current year earnings to enrich the current year picture include the following:

a) Past earnings – Income averaging relies on the permanent income hypothesis to treat career earnings as the only relevant evidence of ability, without regard to the earnings sequence. As the NDPF literature emphasizes, however, information

---

<sup>83</sup> Cites.

about ability levels plays out over time.<sup>84</sup> A high-earning year is informative about a given individual's likely ability level even if subsequent earnings are lower, since the decline, while it might reflect an adverse shock to ability, might also reflect a voluntary change in labor supply. A rational and benign government that was engaged in providing insurance against ability risk would inevitably make use of information about past earnings in deciding what taxes to levy in a given period.<sup>85</sup>

One possible implication is what one might call reverse income averaging.<sup>86</sup> In Vickrey-style cumulative averaging, a taxpayer with high earnings in Year 1 and low earnings in Year 2 will tend to pay less in Year 2 than she would have under a purely annual system, because in effect she gets a refund for Year 1 as the measure of her ability is revised downward. In an NDPF framework, the implication might instead be levying a higher tax in Year 2 than we would have based purely on annual information, because we have evidence suggesting reduced work effort (notwithstanding the possibility that there has instead been an adverse shock to ability).<sup>87</sup> This implication is all the stronger if the taxpayer has significant savings, which could have helped to motivate the decision to work less by providing an alternative source of current financial support.<sup>88</sup>

b) Savings – The NDPF literature closely associates imposing higher taxes by reason of past high earnings with taxing saving. From the perspective of the current

---

<sup>84</sup> See, e.g., Golosov, Kocherlakota, and Tsyvinski, *supra*, at 580.

<sup>85</sup> See William Rogerson, Repeated Moral Hazard, 53 *Econometrica* 69 (1985) (considering the importance of memory in repeated principal-agent interactions with moral hazard); Golosov, Tsyvinski, and Werning, *supra* at 20 (describing the dynamic time inconsistency problem in this framework, where a government that cannot credibly commit against subsequently exploiting early-years labor supply information about ability may overly discourage labor supply in such years).

<sup>86</sup> Jeff Strnad suggested this phrase to me.

<sup>87</sup> See Kevin Roberts, Theoretical Limits to Redistribution, 51 *Rev. Econ. Stud.* 177 (1984);

<sup>88</sup> See Narayana Kocherlakota, Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, 73 *Econometrica* 1587 (2005); Stefania Albanesi and Christopher Sleet, Dynamic Optimal Taxation with Private Information, 73 *Rev. Econ. Stud.* 1 (2006).

period, high saving may help to finance (and thus explain) an otherwise surprisingly high decline in earnings.<sup>89</sup> From the perspective of past periods, since someone planning to reduce her labor supply would be expected to save, observing high saving can in effect serve as a proxy for directly observing high past earnings.<sup>90</sup> This association between conditioning taxes on past earnings and on savings reflects the overlap between issues of income averaging and of tax base choice. However, since the use of savings in NDPF models relates more directly and obviously to the choice of ideal tax base, I defer discussing it to section IV.

c. Age-related wage distribution and labor supply elasticity – A recent article by Michael Kremer<sup>91</sup> makes the point that an optimal tax structure might impose distinct marginal rates on people of different ages, for two reasons. The first is that labor supply elasticity may vary with age. In particular, young and old people tend to have more elastic labor supply than people in mid-career. Second, the annual earnings distribution varies with age, with young and old people being more concentrated at the low end of the scale than people in mid-career. Thus, the rates applying at low and medium income levels will apply at the margin (and thus to the next dollar of potential earnings) to a smaller proportion of mid-career than of young and old workers. Both considerations suggest that young and old people should generally face lower marginal rates than mid-career workers.

These arguments for age-based taxation are intellectually consistent with the permanent income hypothesis, since they do not rebut the claim that one allocates

---

<sup>89</sup> See Kocherlakota, Zero Expected Wealth Taxes, *supra*.

<sup>90</sup> See Albanesi and Slaeet, *supra* at 25 (noting that their model, unlike an alternative one that observes past earnings directly, “uses[s] wealth to summarize aspects of an agent’s past history.”)

<sup>91</sup> Michael Kremer, “Should Taxes Be Independent of Age?”, available at [http://post.economics.harvard.edu/faculty/kremer/webpapers/Should\\_Taxes.wpd](http://post.economics.harvard.edu/faculty/kremer/webpapers/Should_Taxes.wpd).

consumption across one's lifetime budget line without regard to the time sequence of earnings. However, they contradict the implication in support of income averaging that tax liability should depend purely on the discounted value of lifetime earnings. They imply once again that taxes depend on when a given dollar of earnings is realized, in this case reflecting additional information about particular stages within the lifecycle.

F. Conclusions Regarding Income Averaging

The problems with the case for income averaging can be interpreted in a couple of different ways. Strictly speaking, the strong logical case for it falls short. It is not generally and universally desirable, even ignoring issues of administrative cost. One should keep in mind, however, that the presumed alternative, a purely annual system, might fare even worse if scrutinized with comparable rigor.

For a purely annual view to hold, each year would have to be completely separate from all others, in two senses. First, people would have to be unable or unwilling to shift work or consumption between years, whether by saving or borrowing. Second, information from other years would have to be irrelevant – not merely potentially less relevant than current-period information – to the assessment of one's relevant current year attributes, such as wage rate, labor supply elasticity, and utility function. This is not compelling as a picture of the world we live in.

Given the problems with a purely annual view, there remains a significant case in support of income averaging's main effect, which is to mitigate the extra tax burdens on people who have fluctuating rather than relatively constant annual incomes. Some design details arguably supported by the preceding analysis are the following:

1) Given the “great expectations” point, the case for relief is weaker for people who newly enter high-earning stages of their careers, such as by entering a high-wage profession after graduating from school. The U.S. income averaging system that was on the books from 1964 through 1986 attempted roughly to address this problem by denying income averaging to certain individuals who had not been self-supporting throughout the averaging period (which went back four years).<sup>92</sup>

2) Given the existence of other tools, such as Social Security and Medicare, to address inadequate retirement saving, arguably there is no reason to let people who retire claim income averaging benefits with respect to their pre-retirement earnings. Moreover, insofar as the timing of one’s retirement is a matter of discretionary personal choice, the decline in income that results from it does not indicate being worse-off or having suffered a negative shock to ability. The payoff to retirement, in the form of eliminating tax liabilities with respect to the foregone earnings, may already be too high under a purely annual system, given both the revenue externality point and the fact that one presumably is not worse off overall, despite the lost income, by reason of retiring voluntarily.<sup>93</sup> Extending the benefits of income averaging in these circumstances would make the arguably excessive tax benefit of retirement greater still.<sup>94</sup>

3) Beginning a career and ending it through retirement are only two examples of status changes across which the case for income averaging is especially weak. Income averaging might also be denied in cases where there is an identifiable change in, say, the taxpayer’s health status or occupation. For age, while denying income averaging between different life stages (such as youth and middle age) may be desirable, the

---

<sup>92</sup> See Internal Revenue Code section 1303(c)(1) (1984); Schmalbeck, *supra*, at 523.

<sup>93</sup> Cf. [my article on household taxation re. secondary earner’s voluntary non-work.]

<sup>94</sup> Cf. Schmalbeck, *supra*, at 572 (arguing against income averaging for retirees).

problem is harder because any particular point at which one tries to draw a line between periods is likely to be arbitrary.

4) At first glance, extending income averaging to individuals with rising incomes but not to those with falling incomes (in the manner of the old U.S. rules) may appear arbitrary. It may very roughly be supported, however, by the point from the NDPF literature that high past earnings are evidence of high current ability. Even if one does not go so far as to embrace reverse income averaging, denying the positive benefits to individuals with falling incomes may be reasonable as a general rule.

5) Given the importance of current cash flows, it may be desirable to structure income averaging in the Vickrey or cumulative manner, rather than in the U.S. and Wisconsin manner, so that current tax payments tend to rise and fall with current income. Given the liquidity problems that people with declining incomes may face, this is all the more important if income averaging is indeed extended to people with falling as well as rising incomes.

#### IV. IMPLICATIONS FOR THE CHOICE BETWEEN INCOME AND CONSUMPTION TAXATION

##### A. Overlap Between the Cases for Income Averaging and for Consumption Taxation

A core argument of this paper is that the cases for income averaging and for consumption taxation are closely linked. This is not to deny that one could logically support one of these two approaches while opposing the other, even ignoring administrative considerations,<sup>95</sup> given that either might be rationalized in a number of different ways. There are indeed many examples of people who have supported one while opposing the other. William Vickrey, for example, supported both income averaging and income taxation.<sup>96</sup> While one could rightly have criticized him for not appreciating the tension (discussed below) between income taxation and a lifetime perspective, he could equally correctly have responded that supporting a tax on saving does not pre-commit one to believing that people with fluctuating earnings should generally pay more tax than those with level earnings. Likewise, one can favor consumption taxation on grounds that are distinct from adopting a lifetime perspective. Edward McCaffery, for example, believes on ethical grounds distinct from a pure welfare economics framework that consumption rather than saving should be taxed, and on similar ethical grounds favors using an annual rate structure to penalize concentrating high consumption in a single year.<sup>97</sup>

---

<sup>95</sup> [Note that shifting to consumption tax might reduce complexity while adopting income averaging might increase it.]

<sup>96</sup> [Cite to Vickrey book]

<sup>97</sup> See Edward J. McCaffery, FAIR NOT FLAT: HOW TO MAKE THE TAX SYSTEM BETTER AND SIMPLER 87-91 (2002).

Within a consistent welfare economics framework, however, the link between the two issues is hard to sever. Distributionally, both follow from viewing lifetime budget lines as the key attribute in discerning total and marginal utility, a stance that relies on the permanent income hypothesis to establish that all equivalent lifetime budget lines are indeed effectively the same. In terms of efficiency, both rely on the Atkinson-Stiglitz analysis. Objections to the underlying assumptions of the permanent income hypothesis, concerning complete markets and consistent rational choice, weigh similarly against both, as does the case for using additional information to supplement the evidence about ability that lifetime earnings provide.

B. The Distributional Case for Consumption Taxation Under the Permanent Income Hypothesis

From a welfare economics perspective, the core distributional argument for a consumption tax is that lifetime earnings determine one's budget line.<sup>98</sup> Savings decisions merely reflect commodity choice within this budget line as between present and future consumption. Thus, individuals who save more and thus derive greater returns to saving (which an income tax reaches and a consumption tax does not) are not relevantly better-off and should not pay more tax on a lifetime basis.

To illustrate, suppose there are two periods and that the rate of return on resources saved in Period 1 for consumption in Period 2 is 5 percent. George and Hilda both earn \$100 in Period 1, meaning that they can consume \$100 in Period 1, \$105 in Period 2, or some lesser combination of goods in each period. Suppose George chooses to consume \$100 in Period 1 while Hilda chooses to consume \$105 in Period 2. While they are relevantly equal, Hilda would pay more than George under an income tax. Suppose the

---

<sup>98</sup> Cites, e.g., Andrews, Bradford, my Tax Notes piece?, Bankman-Weisbach?

tax rate was 40 percent. Both would pay \$40 in Period 1, while Hilda would pay \$2 and George would pay zero in Period 2. By contrast, under a consumption tax, they would effectively pay the same amount. Given the 5 percent interest rate, George's \$40 tax payment in Period 1 would have the same present value as Hilda's \$42 tax payment in Period 2. Finally, consider an earnings tax, which is equivalent to a consumption tax since returns to saving are excluded from both.<sup>99</sup> Once again George and Hilda would be treated the same, here by reason of paying \$40 each in Period 1 and zero in Period 2.

In this hypothetical, the only required capital market is one permitting saving at a 5 percent rate between Periods 1 and 2. Suppose, however, that George and Hilda had each earned \$105 in Period 2, instead of earning \$100 in Period 1. The analysis would remain the same so long as available capital markets permitted George to borrow \$100 in Period 1 against his expected earnings. Now the mechanism by which an income tax, unlike a consumption tax, would treat him better than Hilda is through the allowance of a \$5 interest deduction in Year 2 when he repays the loan.<sup>100</sup>

A consumption tax, by taxing George and Hilda the same, gets the correct distributional result if they are relevantly identical given that they have the same lifetime incomes. This equivalence between George and Hilda follows from the permanent income hypothesis plus the absence of other pertinent distinctions between them.

Now suppose we reject the equivalence of individuals with the same lifetime incomes, because incomplete markets or departures from consistent rational choice make the exact sequences of earning or consumption important. This does not necessarily

---

<sup>99</sup> Cite, e.g., Bankman-Weisbach.

<sup>100</sup> The actual existing income tax disallows deductions for consumer interest. See Internal Revenue Code section 163(h)(1) and (2). In principal, however, the argument for allowing interest deductions on all dissaving under an income tax is identical to the argument for taxing the positive interest income on net saving. [Cites]

make the case for an income tax, which appears to rest on treating income, including returns to saving that reflect commodity choice rather than different opportunity sets, as synonymous with the normative concept of ability to pay.<sup>101</sup> It does, however, move towards leveling the intellectual playing field between the two tax bases, by weakening a compelling argument for consumption taxation.

What is more, viewing periods as separate rather than as linked is generally more consistent with an income tax than a consumption tax framework. If one is looking purely at consumption opportunities (and thus ability to pay) within a given period, then all of one's wealth is relevant, and an income tax at least comes closer than a consumption tax to achieving the wealth tax ideal by taxing the return to wealth. Moreover, the argument for a consumption tax that, rather than ignoring unspent wealth, it merely defers payment of the tax on such wealth (without any reduction in tax burden) until consumption occurs<sup>102</sup> loses force if we are focused on the current period and do not regard as relevant what might happen in future periods. The classic argument that an income tax better measures ability to pay than does a consumption tax, because it takes account of unspent wealth,<sup>103</sup> is explicitly present period-focused.

C. The Efficiency Case for Consumption Taxation Under the Permanent Income Hypothesis

It has been known for a considerable time that an income tax distorts two choices, by discouraging both work and saving, while a consumption tax distorts only one, by discouraging work.<sup>104</sup> The belief that this creates an efficiency tradeoff between the two

---

<sup>101</sup> Cites.

<sup>102</sup> Cites, e.g., to my Tax Notes article.

<sup>103</sup> Cites.

<sup>104</sup> Cites.

taxes, and that an income tax might be more efficient because its imposing a tax on saving permits it to impose a lower tax burden on work, is still occasionally expressed.<sup>105</sup> However, as Joseph Bankman and David Weisbach recently showed, it is logically incorrect.<sup>106</sup> They note that an income tax is effectively a differential commodity tax, in which future consumption goods are taxed at a higher rate than current consumption goods, violating the Atkinson-Stiglitz theorem. Distorting decisions regarding when to consume, by taxing future commodities at a higher rate than current commodities, does nothing to mitigate the underlying labor supply distortion. Instead, the discouragement of saving for future consumption is simply layered on top, adding to total distortion.<sup>107</sup>

As noted earlier, this is equivalent to arguing for income averaging within a consumption tax, since otherwise the Atkinson-Stiglitz theorem is violated by imposing different marginal rates on consumption in different years. It likewise is akin to the efficiency argument for income averaging that emphasizes rate smoothing as to earnings in different years, except that here the purely additive distortion relates to when one works, rather than to when one consumes. The efficiency case for consumption taxation, like that for income averaging, thereby relies on viewing the entire lifespan as a single uniform period, a view that is easiest to support if we assume complete markets, consistent rational choice, and lack of other pertinent information. As I discuss next, modifying these assumptions muddies the case for consumption taxation, just as it does for income averaging.

#### D. Problems With the Case for Consumption Taxation

##### 1. Incomplete Markets

---

<sup>105</sup> See, e.g., Gruber, *supra*, at 708.

<sup>106</sup> Bankman and Weisbach, *supra*, at \_\_.

<sup>107</sup> See *id.* at \_\_.

As noted in section III, when markets are incomplete, as in the “great expectations” scenario and when earning ability is risky, lifetime comparisons are thrown off, and individuals with the same lifetime incomes may differ in total utility and have varying period-specific marginal utilities even if they are otherwise identical. This undermines the equity case for consumption taxation, albeit possibly without otherwise advancing the case for income taxation,<sup>108</sup> by contradicting reliance on lifetime income as the proper standard of comparison for distributional purposes.

## 2. Departures from Consistent Rational Choice

Departures from consistent rational choice, like incomplete markets, to some extent simply disrupt the clean logical case for consumption taxation without putting anything else in its place, indicating that we do not know enough to make confident claims about total and marginal utility. In one important respect, however, the case for income taxation may be more directly aided. This relates to the possibility that high saving is itself evidence of ability, rather than simply or even primarily evidencing a greater preference for later as compared to current consumption.

Suppose we believe that most people, given their utility functions, should engage in substantial lifetime consumption smoothing, including saving adequately for retirement, but that this requires mental and emotional skills that the population holds very unevenly. High savers might, in this scenario, on average be more patient and farsighted than low savers, having greater self-control and capacity to restrain counter-

---

<sup>108</sup> In a truly comprehensive or Haig-Simons income tax, fluctuations in the value of human capital would have current income tax consequences, just like changes in the value of any other asset. (Cite Kaplow.) This would suggest taxing individuals with “great expectations” on the appreciation of their human capital even if they had no access to the value before the future labor income started being realized. This arguably would be in tension with the underlying normative basis for an income tax, if we think of ability to pay, defined in terms of income, as having something to do with currently accessible resources.

productive impulses. In addition, they might more generally be sufficiently more reflective about their opportunities to do a better job of optimizing their choices, such as by modifying simple rules of thumb where it is feasible to do better.

Under such a view, higher saving clearly would be evidence of a broader ability of some kind. The harder question is what sort of ability, with what sorts of implications for distribution policy. Suppose initially that having these skills implied high earning ability, independently of the evidence offered by observed wages. This might, for example, reflect that savers benefit within the labor market from being more patient and far-sighted than non-savers, enabling them to earn more even if their skills are otherwise the same. Taxing savers more than non-savers who had the same lifetime incomes might then be distributionally optimal, for the same reason as taxing high-earners more than low-earners.<sup>109</sup> Under this view, while an income tax would not necessarily combine the two types of information optimally, at least it would be making positive use of a type of information that a consumption tax, by being savings-neutral, ignores.

Now suppose instead, however, that high saving merely denotes ability as a consumer – that is, the capacity to derive more utility than others could from using the same resources, by properly deploying them to the point in one's lifecycle where they improve wellbeing the most. This type of ability has mixed implications from a welfare economics perspective. On the one hand, the abler consumer presumably has the higher total utility, supporting redistribution to less able consumers if we believe that the marginal utility of a dollar generally declines with rising total utility, and/or if we give independent weight to equality in wellbeing. On the other hand, the fact that the abler

---

<sup>109</sup> Emmanuel Saez, The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes, 83 J. Pub. Econ. 217, 228 (2002)

consumer can derive more utility from an extra dollar suggests redistributing to her, all else equal, rather than away from her.

### 3. Additional Information

An income tax uses more information than does a consumption tax, since it makes lifetime net tax liability depend on savings decisions and thus on the timing of consumption relative to earning. In one important respect, this extra information may improve the performance of the income tax, relative to that of the consumption tax, as a redistributive tool. The NDPF literature strongly suggests that taxing savings, at least when it is at high levels or associated with past high earnings but perhaps more generally, may actually be optimal.<sup>110</sup>

The basic point is that, with high savings, one can more easily afford to under-utilize one's earning ability by working less, thereby in effect camouflaging oneself as a lower-ability individual than one actually is. Decisions to work less impose a negative revenue externality on the government, thus reducing its ability to provide workers with insurance against having low ability, and requiring it to meet any further revenue needs in some other way that may involve increased distortion. This negative effect of savings on labor supply potentially makes it "optimal for society to deter savings by taxing it."<sup>111</sup>

Articles in the NDPF literature or tradition have made this point at least since an important 1978 article by the leading economists Peter Diamond and James Mirrlees.<sup>112</sup>

While some models suggest imposing no net tax on saving – for example, by having

---

<sup>110</sup> This point was apparently first made in Peter Diamond and James Mirrlees, A Model of Social Insurance With Variable Retirement, 10 J. Pub. Econ. 295 (1978). For more recent examples, see Kocherlakota, Wedges and Taxes, 94(2)AEA Papers and Proceedings 109 (2004); Kocherlakota, Zero Expected Wealth Taxes, *supra*; Golosov, Kocherlakota, and Tsyvinski, *supra*.

<sup>111</sup> Golosov, Kocherlakota, and Tsyvinski, *supra*, at 577.

<sup>112</sup> Diamond and Mirrlees, *supra*.

taxes on high savers or those whose earnings surprisingly decline offset by subsidies to low savers or those whose earnings surprisingly increase<sup>113</sup> – others agree with Diamond and Mirrlees in finding a general case for taxing saving in response to the externality. In particular, a recent article by Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski finds a tax on saving to be optimal, by reason of the labor supply effect and consequent revenue externality, in a relatively general setting. These authors conclude that the Atkinson-Stiglitz theorem, while applicable to support the desirability of uniform commodity taxation as to the items that are simultaneously available at any time, does not apply (as recent legal authors such as Joseph Bankman and David Weisbach have argued)<sup>114</sup> to support a zero rate of capital income taxation.<sup>115</sup>

#### E. Conclusions Regarding Consumption Taxation

Departures from the assumptions underlying the permanent income hypothesis and additional information have four effects on the otherwise compelling welfare economics case for consumption taxation. First, they introduce enough noise to make any definite conclusion about the ideal system less tenable than it would otherwise be. Second, by suggesting that current period circumstances may at times be more relevant than information about other periods, they support a more present-focused system than one focused, like a consumption tax, on lifetime income. While this may not directly transfer into support for income taxation, it may add to the distributional relevance for the current period of unconsumed wealth. An income tax, unlike a consumption tax, assigns

---

<sup>113</sup> See Kocherlakota, *Wedges and Taxes*, supra (suggesting tax on high savers and subsidy to low savers if future skills are uncertain); Kocherlakota, *Zero Expected Wealth Taxes*, supra (suggesting positive wealth tax for individuals whose earnings surprisingly fall, along with a wealth subsidy for those whose earnings surprisingly rise).

<sup>114</sup> Bankman and Weisbach, supra.

<sup>115</sup> Golosov, Kocherlakota, and Tsyvinski, supra at 580.

current period tax consequences to unconsumed wealth if there is a positive return to such wealth. Third, if higher saving, all else equal, tends to be a signal of earning ability, then taxing returns to saving may further the core distributional aim of providing social insurance against “ability risk.” (The consequences are more ambiguous, however, insofar as high savings indicate ability as a consumer, in the sense of being able to extract more utility from the same resources.) Finally, taxing saving may be socially optimal given that savings can reduce labor supply, thereby generating a negative revenue externality that adversely affects ability insurance and may require higher distortionary taxes to help finance government purchases.

How strongly do these considerations undermine the recent shift in academic viewpoints towards a new consensus favoring consumption-based rather than income-based tax reform? This is ultimately a matter of judgment about the significance of the factors noted above. One’s judgment also should be affected by administrative and political economy issues concerning how the two rival tax bases are likely to operate in practice. I myself continue to favor replacing the current income tax with a progressive consumption tax, but in large part for a relatively humdrum, rather than abstractly theoretical, reason, which is that, as William Andrews noted [more than] thirty years ago, a realization requirement remains the Achilles heel of any practically feasible income tax.<sup>116</sup> This is not as intellectually satisfying as relying on the implications of the permanent income hypothesis, but it also may prove harder to rebut.

---

<sup>116</sup> Cite Andrews, & something on the problems associated with realization.

## V. CONCLUSION

The permanent income hypothesis is a powerful and appealing idea, logically applying basic precepts of neoclassical economics. It also has the pleasant consequence of making the normative analysis of ideal tax policy a lot simpler, by establishing a compelling case, within the assumptions of welfare economics, both for consumption taxation and for lifetime income averaging.

Unfortunately from the standpoint of making life simple, the premises that the permanent income hypothesis requires do not entirely hold. First, markets are not complete. For example, it may be hard to borrow against one's future expected earnings, and it is impossible to change past consumption decisions by reason of new information about available lifetime resources. Second, people do not always exhibit consistent rational choice across time. For example, they may be myopic or prone to hyperbolic discounting, and they may use mental accounts, in deciding how to use various dollars, that lead them to violate the principle that a dollar is a dollar is a dollar. These departures increase the relative importance of current period information, thereby exposing as overly flat and undifferentiated the one-period picture of an entire lifespan that the lifetime income concept offers.

Two next steps in the tax policy debate seem indicated by the analysis in this article. The first is exploring how income averaging might work, as a technical matter, while limited to the types of circumstances where I have argued that it is most appropriate (i.e., within what is effectively a single period for the taxpayer, subject to internal consumption smoothing and without significant changes in current earning ability). The second is further exploring, in keeping with the insights from the NDPF

literature but attempting further to gauge their real world significance, to what extent the case for income taxation is intellectually rehabilitated, and how other than in a straightforward income tax one might tax saving in circumstances where the argument for doing so is strongest.