

FALL 2021
NEW YORK UNIVERSITY
SCHOOL OF LAW

Law and the New Dynamic Public Finance
Daniel Hemel

September 28, 2021
Vanderbilt Hall – 208
Time: 2:15 – 4:15pm
Accessible via Zoom
Week 1

SCHEDULE FOR FALL 2021 NYU TAX POLICY COLLOQUIUM
(All sessions meet from 2:15 - 4:15 pm in Vanderbilt 208, NYU Law School)

1. Tuesday, September 14 – Jake Brooks and David Gamage, The Indirect Tax Canon, Apportionment, and Drafting a Constitutional Wealth Tax.
2. Tuesday, September 28 – Daniel Hemel, Law and the New Dynamic Public Finance.
3. Tuesday, October 12 – Jennifer Blouin, article TBD.
4. Tuesday, October 26 – Manoj Viswanathan, Retheorizing Progressive Taxation.
5. Tuesday, November 9 – Ruth Mason & Michael Knoll, [Undue Burdens Under the Dormant Commerce Clause]
6. Tuesday, November 23 – Mindy Herzfeld, Taxes Are Not Binary: The Unfortunate Consequences of Splitting Taxes Into Arbitrary Categories.
7. Tuesday, November 30 – Alan Auerbach, Taxes and Low Interest Rates.

Law and the New Dynamic Public Finance

Daniel Hemel*
September 15, 2021

Introduction

Over the last two decades, a new movement in academic economics has challenged conventional wisdoms in optimal tax theory and generated fresh insights for real-world tax policy. Known as “the new dynamic public finance,” this movement has altered the way that economists think about labor income taxation, capital taxation, and the credibility of tax policy over time. Along the way, the NDPF literature has identified new justifications for previously perplexing features of the existing tax-and-transfer system and has called other elements of the status quo into serious question.

Mainstream economics has embraced the new dynamic public finance revolution. All the top peer-reviewed economics journals publish NDPF papers.¹ Undergraduate public finance textbooks cover basic NDPF concepts.² But legal scholars—including scholars of tax law—have largely ignored the emergence of NDPF. One notable exception is Daniel Shaviro, whose 2007 article “Beyond the Pro-Consumption Tax Consensus” highlighted NDPF’s implications for income-averaging proposals and the choice between income and consumption tax bases.³ Since then, though, only seven law review articles in the Westlaw database have even mentioned “the new dynamic public finance,” and none has sought to take stock of NDPF’s wide-ranging implications for legal analysis.

That ought to change. This paper presents the case for “law and the new dynamic public finance” as fertile ground for academic exploration. Like “law and public choice,”⁴ “game theory and the law,”⁵ and “behavioral law and economics,”⁶ law and the new dynamic public finance

* Professor of Law, University of Chicago Law School; Visiting Professor of Law, New York University School of Law (2021-2022 academic year). This paper is a significantly revised and expanded version of the 2021 Coase Lecture at the University of Chicago Law School. For helpful comments on previous drafts, the author thanks Dhammika Dharmapala, Peter Ganong, Jacob Goldin, Richard Schmalbeck, Lawrence Zelenak, and participants in the Duke Law School Tax Policy Workshop. All errors are my own.

¹ See, e.g., Mikhail Golosov, Narayana Kocherlakota & Aleh Tsyvinski, Optimal Indirect and Capital Taxation, 70 Rev. Econ. Stud. 569 (2003); Mikhail Golosov, Maxim Troshkin & Aleh Tsyvinski, Redistribution and Social Insurance, 106 Am. Econ. Rev. 359 (2016); Mikhail Golosov & Aleh Tsyvinski, 114 J. Pol. Econ. 257 (2006); Narayana R. Kocherlakota, Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, 73 Econometrica 1587 (2005); Iván Werning, Optimal Fiscal Policy with Redistribution, 122 Q.J. Econ. 925 (2007).

² [Cite to latest edition of Gruber textbook.]

³ See Daniel Shaviro, Beyond the Pro-Consumption Tax Consensus, 60 Stan. L. Rev. 745 (2007).

⁴ See Daniel A. Farber & Philip P. Frickey, Law and Public Choice (University of Chicago Press, 1991).

⁵ See Douglas G. Baird, Robert H. Gertner & Randal C. Picker, Game Theory and the Law (Harvard University Press, 1998).

⁶ See Christine Jolls, Cass R. Sunstein & Richard Thaler, A Behavioral Approach to Law and Economics, 50 Stan. L. Rev. 1471 (1998).

gives legal scholars new purchase on old questions and suggests new puzzles for further study. It also has the potential to inform and enrich economics research, making “law and the new dynamic public finance” a two-way street.

The new dynamic public finance focuses on how policy mediates the consumption effects of productivity shocks. Its central concerns are (a) changing productivities and (b) changing policies. “Productivity,” often denoted by the Greek letter theta (θ), refers to the rate at which an individual can convert labor effort into output of consumption goods. The policies studied in the NDPF literature are policies that determine the relationship between an individual’s output and her consumption. Income taxation is one such policy, but the NDPF approach is similarly well-suited for the study of other policies that cause individuals to consume more or less than their output (e.g., disability insurance, unemployment compensation, and private-sector insurance and income-pooling arrangements).

Whereas earlier efforts in optimal tax theory characterized each individual’s productivity, or θ , as a stable attribute, the NDPF literature understands θ as a value that evolves stochastically over the life cycle. Shocks to an individual’s θ can be negative or positive. For example, Covid-19 was a large negative shock to the productivity of individuals who contracted the disease and experienced long-lasting and debilitating sequelae. On the other hand, the pandemic was a positive productivity shock for developers of video-conferencing technology (e.g., the founder and top employees of Zoom). The policies studied in the NDPF literature seek to provide insurance against negative θ shocks while also mitigating moral hazard (i.e., the incentive for high- θ types to present themselves as having low θ values so that they can pay less in taxes and/or receive more in transfers).

Earlier work in optimal tax theory assumed that the government either chose a one-time policy to mediate the relationship between output and consumption or that the government could credibly commit to future policies. By contrast, the NDPF literature grapples seriously with the challenge of policy instability. The NDPF literature asks questions such as: How do policies today affect individuals’ expectations about policies in the future? And what policies can governments (or other providers of θ -shock insurance) adopt today that will mitigate the potentially distortionary effects of policy uncertainty?

Much of the NDPF literature is highly mathematical and considers a stylized version of the tax-and-transfer system. Law and the new dynamic public finance maps NDPF insights onto real-world θ -shock insurance schemes. The mapping project is both positive and normative. On the positive side, it clarifies the terms of the θ -shock insurance arrangements embedded in existing policies. On the normative side, it fleshes out the details of ideal θ -shock insurance schemes and suggests reforms that would shift the status quo closer to the optimum. Tax law lies at the center of the project’s attention, but law and the new dynamic public finance also offers lessons and poses questions for the study of contract law, tort law, property law, and constitutional law—all of which affect individuals’ vulnerability to θ shocks, their susceptibility to moral hazard, and/or the ability of governments to credibly commit to future policies.

The law-and-NDPF project bears special relevance to the problem of inequality, which former President Barack Obama has described as “the defining challenge of our time.”⁷ Income inequality is a manifestation of θ shocks—the initial θ shock resulting from the natural lottery and subsequent θ shocks over the course of childhood and adulthood. Law and the new dynamic finance offers a conceptual framework for understanding and evaluating policy responses to income inequality. It also casts further light on inequality’s underappreciated social costs.

[Part I of this paper provides an overview of the NDPF approach. Part II explores NDPF’s implications for tax-and-transfer policy. Part III considers connections to other areas of law—including contracts, torts, property, and constitutional law.]

I. An Introduction to the New Dynamic Public Finance

The name “new dynamic public finance” comes from the title of a lecture by economist Narayana Kocherlakota, later the president of the Federal Reserve Bank of Minneapolis, at a conference in Florence in 2004. “I wanted to choose a title for the talk that would generate attendance and also signal that I wanted to discuss an agenda broader than of my own individual papers,” Kocherlakota writes.⁸ The title he chose was “deliberately catchy.”⁹ Somewhat to Kocherlakota’s surprise, it stuck. The name “new dynamic public finance” is now used to refer to a body of work that applies similar methods to similar questions.¹⁰ Like other movements within economics (e.g., behavioral economics), its contours are not entirely clear, but a core set of insights can be identified and attributed to NDPF. This part introduces those insights.

A. The “Old” Public Finance

To explain NDPF’s contributions, it will be helpful to start by briefly surveying the landscape of optimal tax theory as it existed prior to NDPF’s arrival—focusing first on optimal labor income taxation and then on optimal capital income taxation.

Mirrlees and Rawls. Optimal tax theory’s main focus—dating back to James Mirrlees’s field-defining 1971 article¹¹—is the construction of a welfare-maximizing tax schedule when the government can observe individuals’ labor income but not their productivity (θ) or,

⁷ White House Office of the Press Sec’y, Remarks by the President on Economic Mobility (Dec. 4, 2013), <https://obamawhitehouse.archives.gov/the-press-office/2013/12/04/remarks-president-economic-mobility>.

⁸ Narayana R. Kocherlakota, *The New Dynamic Public Finance* 2 n.1 (Princeton University Press, 2010).

⁹ *Id.*

¹⁰ The body of work begins with P.A. Diamond & J.A. Mirrlees, *A Model of Social Insurance with Variable Retirement*, 10 *Journal of Public Economics* 295 (1978). The Diamond-Mirrlees paper is definitely dynamic—though not that new. NDPF is sometimes described as the “dynamic Mirrleesian” approach in light of Mirrlees’s extraordinary influence over the field. See, e.g., Stefanie Stantcheva, *Dynamic Taxation*, 12 *Annual Rev. Econ.* 801 (2020).

¹¹ See J. A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38 *Rev. Econ. Stud.* 175 (1971).

equivalently, their labor effort.¹² Mirrlees’s 1971 article considered several different distributions of θ , but for expositional purposes, it will be convenient to imagine θ as binary (i.e., either “high” or “low”). It will also be helpful to imagine only two possible wage levels: “high” and “low.” High- θ types can choose high or low wage levels, whereas low- θ individuals are consigned to low wage levels.¹³ We will assume that high- θ types who choose low wage levels enjoy a high amount of leisure, while high- θ types who choose high wage levels and low- θ types who choose low wage levels enjoy less leisure.

Table 1. Binary θ Types and Wage Levels—Single Period

	High Wage	Low Wage
High- θ	<u>Consumption</u> : High wage less tax <u>Leisure</u> : Low	<u>Consumption</u> : Low wage plus transfer <u>Leisure</u> : High
Low- θ		<u>Consumption</u> : Low wage plus transfer <u>Leisure</u> : Low

In optimal tax models, a tax schedule is chosen to maximize social welfare. This requires some specification of a social welfare function. For present purposes, let’s assume that the social welfare function accords dispositive weight to the consumption of low- θ types. Such a welfare function might be rooted in utilitarianism (e.g., if the marginal utility of consumption for high-wage earners is essentially zero), or it might be based on a form of prioritarianism (which accords greater weight to the welfare of the worse off).

The optimal tax schedule under these circumstances sets the tax on high wage earners at the revenue-maximizing level. In other words, the rate on high wage earners is as high as it can go without inducing high- θ types to “mimic” low- θ types (i.e., to choose low wage levels, which for them means high leisure). When the tax rate on high wage earners is optimal, any further increase will make low- θ types worse off by leaving the government with less revenue to redistribute. The tax rate on high wage earners is thus subject to an incentive compatibility constraint: any change that makes high wage earners worse off must be offset by a change that makes high wage earners better off or else high- θ types will choose to mimic.

A different way of framing the central problem in optimal tax theory borrows from another 1971 publication, one that is more familiar to most non-tax lawyers than Mirrlees’s article: John Rawls’s *A Theory of Justice*.¹⁴ Rawls asks us to imagine ourselves in the “original position,” behind a “veil of ignorance” that prevents us from knowing the circumstances of our lives.¹⁵ In other words: If we could purchase θ -shock insurance before we knew the outcome of the initial θ shock (the natural lottery), what would we want the terms of that insurance scheme to be? Setting the tax on high wage earners at the revenue-maximizing rate and

¹² “Equivalently” because if the government can observe labor income (y) and labor effort (l), it can infer productivity: $\theta = y/l$.

¹³ This two-part structure abstracts away from issues such as the extensive versus intensive margins of labor supply.

¹⁴ John Rawls, *A Theory of Justice* (Harvard University Press, 1971).

¹⁵ *Id.* at 136-142.

redistributing the revenue to low wage earners accords with Rawls's maximin principle, which ranks alternatives based on the outcomes they deliver for the worst-off members of society.¹⁶

The Atkinson-Stiglitz Theorem. An influential 1976 article by Tony Atkinson and Joseph Stiglitz extends optimal tax theory to an environment with two periods and the possibility of saving.¹⁷ Now, individuals work in period 1 and retire in period 2. They can save some of their period 1 earnings for period 2 consumption. The question considered by Atkinson and Stiglitz is whether it is optimal to impose any additional tax on capital income (i.e., income generated by saving). Atkinson and Stiglitz's answer, in a nutshell, is no: the government should allow for tax-free saving.

Atkinson and Stiglitz draw an analogy between a capital income tax and a nonuniform commodity tax. They characterize period 1 consumption and period 2 consumption as different goods (or commodities) that can both be purchased with period 1 labor income. If the interest rate is 10 percent and the capital income tax rate is zero, then an individual can substitute \$1 of period 1 consumption for \$1.10 of period 2 consumption. Individuals will allocate their period 1 labor income between period 1 and period 2 consumption until, at the margin, they are indifferent between an extra \$1 of consumption in period 1 and an extra \$1.10 of consumption in period 2.

Within this framework, capital income taxation produces two effects. First, it distorts individuals' choices between period 1 and period 2 consumption. For example, a 20 percent capital income tax means that \$1 of period 1 consumption can be substituted for \$1.08 of period 2 consumption (rather than \$1.10). The change in the intertemporal rate of substitution (i.e., the rate at which period 1 consumption can be substituted for period 2 consumption) will affect the consumption bundle that individuals choose. Conditional on an assumption discussed in the margin,¹⁸ the capital income tax will cause individuals to allocate a larger portion of their period 1 labor income to period 1 consumption than to now-more-expensive period 2 consumption.

Second, and more subtly, capital income taxation distorts individuals' choices between period 1 leisure and period 1 labor. Assuming that individuals allocate at least some of their period 1 labor income to period 2 consumption, the capital income tax reduces the value of the

¹⁶ Id. at 152-53.

¹⁷ A.B. Atkinson & J.E. Stiglitz, *The Design of Tax Structure: Direct versus Indirect Taxation*, 6 *J. Pub. Econ.* 55 (1976). For a fuller explanation of the Atkinson-Stiglitz model that is accessible to readers with no economics training, see Joseph Bankman & David A. Weisbach, *The Superiority of an Ideal Consumption Tax over an Ideal Income Tax*, 58 *Stan. L. Rev.* 1413 (2006).

¹⁸ The statement in text assumes that the intertemporal elasticity of substitution is positive—in other words, the substitution effect dominates the income effect. In theory, the intertemporal elasticity of substitution could be negative (e.g., individuals could be “target savers” who allocate a larger fraction of their labor income to period 2 consumption when period 2 consumption becomes more expensive). However, the bulk of the empirical evidence suggests a positive intertemporal elasticity of substitution. See generally Tomas Havranek, Roman Horvath, Zuzana Irsova & Marek Rusnak, *Cross-Country Heterogeneity in Intertemporal Substitution*, 96 *J. Int'l Econ.* 100 (2015) (mean value of 0.5 in meta-analysis of 2735 estimates from 104 countries).

consumption bundle that individuals can purchase with period 1 labor income. This, in turn, makes period 1 leisure relatively more attractive than it was before.

To see why these two distortions are suboptimal, imagine what would happen if we started with the optimal period 1 labor income tax and added a period 2 capital income tax. Recall that we are already, by assumption, taxing high wage earners as much as possible without inducing them to mimic. A period 2 capital income tax, by further reducing the bundle of consumption associated with high wage levels, would therefore induce high- θ individuals to earn low wages in period 1. The incentive compatibility constraint binds: starting from an optimal period 1 labor income tax, any period 2 capital income tax must be accompanied by a cut in the period 1 labor income tax such that high- θ types still choose high period 1 wage levels.

But consider the following: For any scenario with positive labor and capital income taxes in which high- θ types still choose high wage levels, there exists an alternative scenario in which those high- θ types would be willing to pay more tax, at least in present value terms, in exchange for eliminating the savings distortion. The opportunity to choose one's own preferred allocation of period 1 and period 2 consumption is itself valuable, which suggests that high wage earners should be willing to pay for that opportunity. The presence of a capital income tax means that the government is not extracting as much as it could from high wage earners: it is not "selling" them the opportunity to choose their preferred intertemporal allocation. Therefore, the presence of a capital income tax tells us that there is an alternative policy that would allow for more redistribution and leave low- θ types better off.

Another way to frame the Atkinson-Stiglitz intuition is to remember why the government is taxing income in the first place: because the government wants to redistribute from high- θ types to low- θ types but cannot observe θ . The government taxes labor income because labor income provides information about θ : the fact that individuals earn high wages means that they are high- θ (though the inverse is not true—the fact that individuals earn low wages does not necessarily mean that they are low- θ). Unless higher- θ types choose to save more than lower- θ types *conditional on period 1 labor income*,¹⁹ an additional tax on capital income simply generates an additional distortion without providing the government with any additional information about anyone's θ .

Relationship to Consumption Taxation. Atkinson and Stiglitz's prescription for tax-free saving can be implemented in two ways. One is to exclude capital income from tax. The other is to tax all income but to allow an immediate deduction for savings. The immediate deduction-

¹⁹ Emmanuel Saez has suggested that higher- θ types do save more than lower- θ types with the same labor income. See Emmanuel Saez, The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes, *J. Pub. Econ.*, 83, 2002, 217-230 ("[T]here is a strong presumption that higher income individuals save more not only because they have more income to save but also because they might have a better financial education and be more aware of the need to save for retirement."). For a rejoinder, see Bankman & Weisbach, The Superiority of an Ideal Consumption Tax over an Ideal Income Tax, 58 *Stan. L. Rev.* 1413, 144-45 (2006).

yield exemption equivalence—sometimes known as the Cary Brown theorem—connects the idea of tax-free saving with consumption taxation.

To illustrate: Imagine that an individual earns \$100 in period 1 and that savings will grow at a 50 percent rate between period 1 and period 2. The individual seeks to smooth consumption across the two periods (i.e., to consume the same amount in each). Now consider two possible tax regimes. In the first, labor income is taxed at a 50 percent rate, and capital income (i.e., the return from savings) is untaxed. The individual will therefore want to save \$20 in period 1, leaving \$50 to pay the period 1 tax and \$30 for period 1 consumption. The \$20 in period 1 savings will grow at a 50 percent rate to \$30 in period 2. The growth will not be subject to tax, since there is no tax on capital income, and thus the individual can consume \$30 again in period 2. (Note that this setup—no deduction for savings but no taxation of capital income—aligns with Roth-style retirement accounts.)

Second, consider the same scenario as above with the following change: Instead of an exemption for capital income, there is a deduction for net savings. This time, the individual will want to save \$40. Thus, her income less net savings in period 1 will be \$100 minus \$40 equals \$60, and her period 1 tax will be \$30. Her \$40 of savings and \$30 of tax will leave \$30 for period 1 consumption. The \$40 of period 1 savings will grow at a 50 percent rate to \$60 in period 2. When she withdraws that \$60 for consumption, she will pay a 50 percent tax (\$30), leaving \$30 for consumption in period 2. (Note that this setup—an immediate deduction for savings with full taxation of dissavings—aligns with traditional-style retirement accounts.)

Table 2. No Capital Taxation (Yield Exemption) vs. Consumption Taxation (Immediate Deduction)

	No Capital Income Taxation (Yield Exemption)	Consumption Taxation (Immediate Deduction)
Period 1		
Labor income	\$100	\$100
Savings	\$20	\$40
Tax	50% x \$100 = \$50	50% x (\$100 – \$40) = \$30
Consumption	\$30	\$30
Period 2		
Savings + growth (1.5x)	\$30	\$60
Tax	\$0	\$30
Consumption	\$30	\$30

As Table 2 and the discussion in text illustrate, a regime of no capital income taxation (the Roth approach) puts the individual in the same position as a tax on income minus net savings (the traditional approach). Either way, she consumes \$30 in each period. The difference between income and net savings is, definitionally, consumption. Thus, we can (at least for now) understand a consumption tax as equivalent to a zero rate of tax on capital income and can understand Atkinson and Stiglitz as arguing for a consumption tax.

Relationship to Chamley-Judd. The Atkinson-Stiglitz zero-capital-tax result shares similarities with another staple of last-quarter-of-the-20th-century public finance theory: the Chamley-Judd result. In the mid-1980s, economists Christophe Chamley and Kenneth Judd both published papers suggesting that the optimal tax rate on capital income is zero in the long run.²⁰ Like the Atkinson-Stiglitz theorem, the Chamley-Judd result forms an important part of the intellectual background to NDPF.

To understand Chamley and Judd's insight, consider again Atkinson and Stiglitz's observation that a capital income tax is analogous to an additional commodity tax on future consumption. With a 20 percent capital income tax rate and a 10 percent interest rate, the additional commodity tax one period out is relatively small: instead of being able to substitute \$1 of period 1 consumption for \$1.10 of period 2 consumption, an individual can substitute \$1 of period 1 consumption for \$1.08 of period 2 consumption. Put differently, the capital income tax is equivalent to a commodity tax on period 2 consumption of 1.8 percent ($\$0.02/\$1.10 \approx 1.8\%$). But now consider how a 20 percent capital income tax with a 10 percent interest rate affects the tradeoff between consumption today and consumption 200 years from now. Absent capital income taxation, \$1 of present-period consumption can be substituted for \$189,905,276 of consumption in 200 years ($\$1 \times 1.10^{200} \approx \$189,905,276$). With a 20 percent capital income tax, \$1 of present-period consumption can be substituted for only \$4,838,950 of consumption two centuries from now ($\$1 \times 1.08^{200} \approx \$4,838,950$). The 20 percent capital income tax is equivalent to a commodity tax on 200-years-from-now consumption of 97.5 percent!

Chamley and Judd conclude that capital income taxation is therefore undesirable in the long run when agents make tradeoffs between current consumption and consumption far off in the future (e.g., a dynastic family planning consumption across generations). Even at relatively low rates, positive capital income taxes translate into very large commodity taxes on far-in-the-future consumption—taxes that fall on the “wrong” side of the Laffer curve. In the long run, capital income taxes lead to capital decumulation as dynasties shift consumption from later periods to earlier periods. And in Chamley and Judd's models, extreme capital decumulation harms not only capitalists but also workers because the disappearance of capital reduces the productivity of labor.

B. The New Dynamic Public Finance Revolution

The new dynamic public finance enriches the analysis of optimal tax schedules by adding the possibility of changing productivities and changing policies. We will consider each in turn.

1. Changing Productivities

²⁰ Christophe Chamley, Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives, 54 *Econometrica* 607 (1986); Kenneth L. Judd, Redistributive Taxation in a Simple Perfect Foresight Model, 28 *J. Pub. Econ.* 59 (1985).

NDPF’s first innovation is to allow for changing productivities across multiple periods of labor. For ease of exposition, we will imagine two periods, again with binary θ values and wage levels. We will also consider the possibility that individuals who have high- θ values in period 1 may subsequently experience negative θ shocks, rendering them low- θ in period 2.

Assume that the tax rate on high wage earners in period 1 is set as high as it can go without inducing high- θ types to mimic. In period 2, some of those period 1 high- θ types may be low wage earners. These “high-lows” (highlighted in Table 3) may be individuals who experienced negative θ shocks and are now truly low- θ . Or they may be individuals who remain high- θ but have chosen a low wage level in order to enjoy more leisure.

Table 3. Binary θ Types and Wage Levels—Two Periods with Negative θ Shocks

	Period 2 High Wage	Period 2 Low Wage
Period 1 High Wage	Period 1: High- θ /Period 2: High- θ “high-high”	Period 1: High- θ /Period 2: High- θ “high-low” (mimicking) Period 1: High- θ /Period 2: Low- θ “high-low” (negative θ shock)
Period 1 Low Wage		Period 1: Low- θ /Period 2: Low- θ “low-low”

History Dependence. One of the core questions in the NDPF literature is whether the tax-and-transfer system should be “history dependent.” Should high-lows receive the same second-period transfer as low-lows (in which case the system would be neutral with respect to past income)? Or should high-lows receive a larger second-period transfer than the low-lows (in which case the system would be “regressive” with respect to past income)? Or, finally, should the high-lows receive a smaller transfer than low-lows (in which case the system would be “progressive” with respect to past income)?

Consider first the case for past-income regressivity (i.e., a larger transfer for high-lows than for low-lows). Increasing the transfer to high-lows will ease the incentive compatibility constraint in period 1: high- θ types in period 1 will have a stronger incentive to earn high wages in period 1 because now, a high period 1 wage guarantees a larger transfer in the event of a negative θ shock in period 2. Because high- θ types in period 1 will assign value to greater protection from negative period 2 θ shocks, they should be willing to pay extra for the additional insurance. Therefore, the period 1 tax rate on high wage earners can rise (relative to the previous revenue-maximizing rate) without inducing high- θ types to mimic in period 1. Even low-lows will be better off because now the government can extract more revenue from high- θ types in period 1: assuming high- θ types are risk averse, they should be willing to pay more for the insurance (at least in present value terms) than what it costs the government to provide.

But while a larger transfer to high-lows eases the incentive compatibility constraint in period 1, it exacerbates the incentive compatibility problem in period 2. Now, earning a low wage in period 2 is more attractive to high- θ types because the transfer to high-lows is larger. In other words, past-income regressivity improves incentives for high- θ types in period 1 but

harms incentives for high- θ types in period 2. And past-income progressivity—i.e., a smaller transfer to high-lows than to low-lows—does the opposite.

As a matter of theory, it is not obvious which of these two approaches is optimal: past-income regressivity or past-income progressivity. Some researchers have tried to make progress on the question by constructing quantitative models of the U.S. economy and then deriving optimal history-dependent tax schedules. In this vein, a recent paper by Marek Kapicka finds that the optimal history-dependent tax system “is more progressive with respect to current income than a history independent tax system, but regressive with respect to past incomes.”²¹ In other words, high-lows ought to pay less in period 2 tax (or receive larger period 2 transfers) than low-lows, which in turn allows for higher period 1 taxes on high wage earners. However, Kacpika acknowledges that the results are quite sensitive to the model inputs.²² NDPF techniques do not definitely resolve the debate between past-income regressivity and past-income progressivity.

While the NDPF approach cannot provide a one-size-fits-all answer to the shape of the optimal history-dependent tax schedule, it can offer comparative statics. Past-income regressivity (i.e., larger transfers to high-lows than to low-lows) will be more desirable when the government (or other θ -shock insurer) is better able to determine whether a high- θ type has experienced a negative θ shock. For example, if the government can observe which individuals experience debilitating injuries that render them work-incapable in period 2, then the government will want to implement a period 2 transfer—conditional on disability—that is regressive with respect to past income. Past-income regressivity in period 2 will ease the incentive compatibility constraint for high- θ types in period 1 without undermining incentives in period 2 (since, by assumption, high- θ types who have not experienced the negative θ shock can’t access the disability-contingent transfer). The worse the government is at verifying negative θ shocks, the weaker the case for past-income regressivity.

Capital Taxation. Another approach to the problem of multiperiod θ -shock insurance emphasizes the role of saving. The “high-low” strategy (i.e., earning a high wage in period 1 and a low wage in period 2) will be more attractive to high- θ types if they can save some of their period 1 income and smooth consumption across the two periods. The government can make the high-low strategy less attractive by taking away the possibility of tax-free saving. The NDPF literature thus suggests a role for capital income taxation: by making the high-low strategy less attractive for high- θ types who do not experience a negative θ shock, a capital income tax can ease the incentive compatibility constraint in period 2.

Imposing a capital income tax in period 2, though, poses the same problem as a capital income tax in the Atkinson-Stiglitz model: it reduces the value of the consumption bundle available to high- θ types who earn high wages in period 1. Whether or not they plan to mimic in

²¹ See Marek Kapicka, Quantifying the Welfare Gains from History Dependent Income Taxation 3 (CEERGE-EI, Working Paper, Jan. 15, 2020).

²² Id. at 38-43.

period 2, high- θ types who earn high wages in period 1 may wish to save some of their income (e.g., because they are still worried about a negative θ shock, or because they have a bequest motive, or because they anticipate a period 3 of retirement). A capital income tax—by reducing the amount of future consumption that can be purchased with a high period 1 wage—makes it less attractive for high- θ types in period 1 to earn high wages (i.e., makes it more attractive for high- θ types in period 1 to mimic). Capital income taxation thus encounters the same tradeoff as history-dependent taxation: changes that improve incentives for high- θ types in period 2 come at the expense of incentives in period 1.

One potential solution is to target the capital income tax at high-lows. In other words, the capital income tax can be regressive over current labor income. Recall that the rationale for the capital income tax is to dissuade high- θ types from presenting themselves as low-wage in period 2. The rationale for imposing the tax therefore does not apply to high-highs. Exempting high-highs from capital income taxation means that (a) high- θ types will have stronger incentives to earn high wages in period 1 because their high period 1 wages will buy more future consumption, and (b) high- θ types will have stronger incentives to earn high wages in period 2 because high period 2 wages will relieve their capital tax burdens.²³

What about high-lows who actually do experience negative θ shocks in period 2? Perhaps counterintuitively, a capital income tax that is regressive over labor income can leave these genuine high-lows better off. High- θ types who don't plan on mimicking in period 2 will presumably save less than those who do—therefore, genuine high-lows will be less affected by the capital income tax than mimicking high-lows. And because the capital income tax on high-lows in period 2 eases the incentive compatibility constraint, the government can provide a larger period 2 transfer to low wage earners.

Notably, the role for capital taxation envisioned by NDPF—detering high- θ types from mimicking low- θ types in period 2—arises only because of the occurrence of negative productivity shocks. If productivity remained constant over the lifecycle, then the government would know that every high-low is mimicking and would want to tax them the same as high-highs. The NDPF rationale for capital taxation responds to the fact that some high-lows *aren't* mimicking. It is the desire to insure period 1 high- θ types against the genuine risk of a negative productivity shock that motivates the capital income tax.

In sum, the NDPF literature points to a role for capital income taxation when individuals are subject to negative productivity shocks. Under those circumstances, a capital income tax can ease the period 2 incentive compatibility constraint by deterring high- θ types from mimicking. Detering high- θ mimickers allows the government to provide larger transfers to low wage earners in period 2, including low wage earners in period 2 who earned high wages in

²³ Indeed, Narayana Kocherlakota—in one of the foundational articles in the NDPF literature—suggests that the capital income tax rate on high-highs should potentially be negative (i.e., high-highs should receive a capital subsidy). See Narayana R. Kocherlakota, Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, 73 *Econometrica* 1587 (2005).

period 1. Thus the capital income tax—especially when targeted at high-lows—facilitates provision of a more complete θ -shock insurance product.

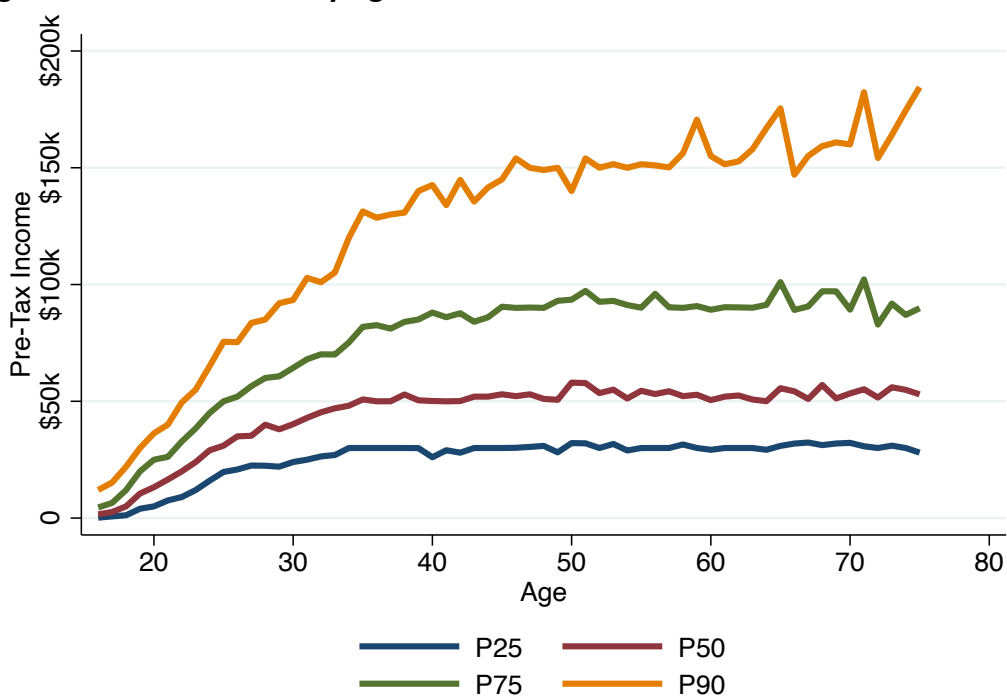
The Role of Age. Age plays a particularly important role in NDPF analyses of labor and capital income taxation. This is so for three reasons. First, as Figure 1 illustrates, income increases over the lifecycle (and especially over the course of the 20s and 30s). If capital markets are perfect, individuals in their 20s and 30s will be able to borrow against their future income in order to smooth consumption. However, if capital markets are imperfect (e.g., because loans with income-contingent repayment plans are subject to adverse selection and moral hazard problems), consumption smoothing may fail, and individuals may “under-consume” in early adulthood (relative to the consumption patterns they would choose with perfect capital markets).²⁴

Second, and even more starkly, inequality increases over the lifecycle. If we focus only on workers in their 20s, the United States appears to be quite egalitarian (though also quite poor). In older age ranges, we evolve into a richer—but much more unequal—society. We can see this phenomenon play out even among elite law school graduates. The income gap between, say, a first-year Cravath, Swaine & Moore associate and a first-year public defender in New York is significant (the Cravath associate makes roughly four times the public defender), but the income gap between a Cravath partner and an experienced public defender in New York is far larger (the Cravath partner makes roughly 60 times the public defender).²⁵

²⁴ See Daniel Shaviro, Permanent Income and the Annual Income Tax (N.Y.U. Law & Econ. Research Paper Series, Working Paper No. 06-28, July 2006), <http://ssrn.com/abstract=920622>.

²⁵ See Average Public Defender Salary: New York, Payscale, https://www.payscale.com/research/US/Job=Public_Defender/Salary (last visited Sept. 14, 2021) (\$48,967 salary for entry-level public defender and \$77,336 salary for experienced public defender); Staci Zaretsky, Cravath Breaks Silence on Associate Raises, Above the Law (June 16, 2021), <https://abovethelaw.com/2021/06/cravath-raises> (\$205,000 salary for Class of 2020 attorneys starting July 1, 2021); The 2021 Am Law 100: Ranked by Profits per Equity Partner, American Lawyer (Apr. 20, 2021), <https://www.law.com/americanlawyer/2021/04/20/the-2021-am-law-100-ranked-by-profits-per-equity-partner> (profits per equity partner of \$4,576,000 reported in 2021). I do not vouch for the accuracy of these salary statistics—the comparison is purely illustrative.

Figure 1. Pre-Tax Income by Age and Percentile



Notes: Current Population Survey (2020) and author's calculations

Third, moral hazard increases over the lifecycle. Relatively few high- θ 55-year-olds leave the labor market in order to enjoy more leisure. Lots of high- θ 65-year-olds do. One recent study estimates that among males, the extensive margin labor supply elasticity in the age 61-65 group is 12 times as high as in the age 51-60 group.²⁶ Thus, the cost of providing θ -shock insurance is much higher for individuals in their 60s than individuals in their 30s, 40s, and 50s. (The extensive margin labor supply elasticity is also very high for individuals in their 20s—and particularly for women in that age demographic.²⁷)

The combination of these three factors suggests that optimal tax schedules are likely to vary by age. First, in the absence of perfect capital markets, we would want the government to help us smooth consumption over the lifecycle by shifting resources to earlier years. One way to do this would be to provide age-dependent lump-sum taxes, with negative taxes (subsidies) for younger individuals and positive taxes for older individuals. The lump-sum taxes would be on top of an income-dependent tax schedule, so very high-income 20-somethings still would pay positive taxes and very low-income 50-somethings still would receive net transfers.

Second, from behind the veil of ignorance, we would want to buy more θ -shock insurance for our 50s than for our 20s and 30s because we face much greater uncertainty about

²⁶ See Marios Karabarbounis, A Road Map for Efficiently Taxing Heterogeneous Agents, 8 Am. J. Econ.: Macroecon. 182, 200 tbl.3 (2016) (Frisch elasticity of labor supply for males estimated to be 1.68 in age 61-65 group and 0.14 in age 51-60). For females, the difference is stark but not quite as stark. See id. (Frisch elasticity of labor supply for females estimated to be 0.81 in age 61-65 group and 0.18 in age 51-60).

²⁷ See id. (Frisch elasticity of labor supply estimated to be 3.94 for females in the age 21-30 group).

income in our 50s than in our 20s and 30s. Income taxes are a form of θ -shock insurance, so this observation suggests that tax rates should rise from early adulthood into one's 50s.

Third, we would likely want to buy less θ -shock insurance for our 60s than for our 50s. This is because θ -shock insurance is costlier when moral hazard is high. The combination of the second and third factors suggests that optimal age-dependent labor income tax rates are "hump shaped." Optimal tax rates rise until at least age 50 as a result of rising intra-cohort inequality and then fall slightly in the 60s as a result of rising moral hazard.²⁸

If tax rates cannot be conditioned explicitly on age, then age-independent capital subsidies can partially implement the desired age-dependent labor income tax schedule. The rationale for capital subsidies mirrors the Atkinson-Stiglitz theorem and the Chamley-Judd result. A capital subsidy allows an individual to buy more future consumption for the same amount of current labor income, and thus makes current-period labor relatively more attractive than current-period leisure. The longer an individual's time horizon, the larger the effect. Imagine, for example, that the retirement age is 65, the interest rate is 10 percent, and the capital subsidy is 1 percent per year. For a 35-year-old, the capital subsidy is equivalent to a -31 percent commodity tax on consumption at retirement. For a 55-year-old, the capital subsidy is equivalent to only a -9 percent commodity tax on consumption at retirement.²⁹ The capital subsidy allows both the 35-year-old and the 55-year-old to purchase larger consumption bundles with their labor, but the effect is more than three times as large for the 35-year-old. The capital subsidy thus emulates an age-dependent labor income tax cut, with a deeper cut for younger workers.

There is, to be sure, no particular reason to assume that tax rates must be age-invariant. The reason we assume that the government cannot condition taxes and transfers on θ is that θ is unobservable. But age is easy to observe and difficult to manipulate, and some tax provisions (e.g., the child tax credit,³⁰ the earned income tax credit,³¹ and the additional standard deduction for the aged³²) already depend on age. One of NDPF's insights is that when we impose age-invariant taxes, we are throwing out a huge amount of information regarding individuals' θ values and vulnerabilities to moral hazard that could be used to provide superior θ -shock insurance products.

2. Changing Policies

The NDPF literature's second major move is to incorporate the possibility that policies might change. Anyone who lived through the early 1990s knows that "read my lips, no new taxes" does not necessarily mean that there will be no new taxes. The risk that the government

²⁸ See Karabarbounis, *supra* note 26, at 207-08 & figs. 6 & 7; Abdoulaye Ndiaye, *Flexible Retirement and Optimal Taxation* (N.Y.U. Stern Sch. of Bus., Working Paper, Aug. 14, 2020).

²⁹ For the 35-year-old: $1.11^{30}/1.1^{30} \approx 1.31$. For the 55-year-old: $1.11^{10}/1.1^{10} \approx 1.09$.

³⁰ I.R.C. § 24(c)(1), (i)(2)-(3).

³¹ I.R.C. § 32(c)(1)(ii)(II).

³² I.R.C. § 63(c)(3).

will renege on its tax commitments has the potential to impose a shadow tax on labor and capital. A common challenge for governments is to persuade their populations that promises about future policy are credible.

The problem is clearest in the capital tax context. A large one-time capital tax would be a very efficient way of redistributing resources from high- θ types to low- θ types.³³ If it were truly one-time, then it would not distort future savings and capital investment decisions. And it might actually encourage labor effort insofar as it would interfere with consumption smoothing by high- θ types who had planned to play the high-low strategy. The potential problem with a large one-time tax on capital is that individuals might expect it to happen again. Fear of confiscation then might induce high- θ types to supply less labor and to decumulate capital.

The NDPF literature offers three responses. The first is to observe that non-taxation of capital today does not necessarily assure individuals that capital taxes will remain low in the near future. Indeed, as the late Emmanuel Farhi and coauthors suggested, non-taxation of capital today can *increase* the risk of confiscation.³⁴ Farhi et al. posit that a government will incur costs when it imposes a confiscatory capital tax—both the fixed costs of enacting and implementing the tax and the (domestic and/or international) reputational costs. The costs are worth bearing only if the benefits are large, and the benefits are larger when wealth inequality is high. By adopting a progressive capital tax, the government can ensure that wealth inequality never exceeds the threshold at which confiscation would become ex-post optimal.

An implication of Farhi et al.'s analysis is that once dynamic consistency considerations come into play, the equity-efficiency tradeoff may not be a tradeoff at all. Greater equity reduces the present-period labor and savings distortions resulting from the shadow of future expropriation. One might view this as the formal version of FDR's intuition that a robust government response to the New Deal could prevent the rise of socialism in the United States.³⁵ Ultimately, a capital income tax may be in capital's interest because the tax reduces the expropriation threat that capital faces.

Second, and cutting in the opposite direction, the NDPF literature suggests that governments can bolster the credibility of their non-confiscation commitments by effectively disabling themselves from taxing capital. Alberto Bisin and Adriano Rampini propose that governments can accomplish this hands-tying objective by giving their citizens access to anonymous financial markets.³⁶ Anonymity reflects a tradeoff: governments lose the ability to use capital taxation in order to improve θ shock insurance, but at the same time, they avoid the large labor and savings distortions that arise when individuals anticipate confiscation. Which

³³ Chamley and Judd both acknowledge this point. See Chamley, *supra* note 20, at 619; Judd, *supra* note 20, at 60.

³⁴ Emmanuel Farhi, Christopher Sleet, Iván Werning & Sevin Yeltekin, Non-linear Capital Taxation Without Commitment, 79 *Review of Economic Studies* 1469 (2012).

³⁵ See Seymour Martin Lipset & Gary Marks, *It Didn't Happen Here: Why Socialism Failed in the United States* (2001).

³⁶ Alberto Bisin & Adriano Rampini, Markets as Beneficial Constraints on the Government, 90 *Journal of Political Economy* 601 (2006).

way the balance ultimately tips may vary from country to country, depending in part on the ability of the country's legal institutions to generate credible commitment through other mechanisms.

Third, the NDPF literature highlights the relationship between government credibility and labor incentives. Lack of credible commitment distorts labor incentives not only because it generates a shadow capital tax, but also because the government may exploit information about past labor incomes. Imagine, for example, that the government announces tomorrow that everyone's tax liability going forward will be based on an age-adjusted five-year average of their income from 2016 to 2020. Nothing we do in 2021 or thereafter will affect our tax liability—in that sense, the tax would be as-if lump sum. The retrospective tax would be a highly efficient way of redistributing from individuals based on their existing θ type. Past income isn't a perfect proxy for θ (indeed, one of NDPF's key insights is that θ evolves over the lifecycle), but it is a pretty good proxy, and the government could therefore accomplish a very large amount of redistribution without any distortion. Yet if individuals anticipate that the government will impose a retrospective labor income tax in the future, then—as in the capital confiscation case—the distortions in the run-up would be severe.

The shadow of retrospectivity connects capital and labor taxation in two ways. First, Farhi et al.'s result—that progressive capital taxation can enhance efficiency by reducing the risk of confiscation—potentially applies to labor too. Retrospective labor income taxation will likely be less attractive to governments when the distribution of labor income—pre- and post-tax—is less lopsided. Second, Bisin and Rampini's observation regarding anonymous financial markets also may apply to anonymous labor markets. (This point will become particularly relevant in the discussion below of value added taxes.)

II. Implications for Taxes and Transfers

While the NDPF literature largely abstracts away from institutional details, NDPF insights yield concrete implications for real-world tax-and-transfer institutions.

History Dependence. Although the federal income tax is not (for the most part) history dependent, other aspects of the U.S. tax-and-transfer system are heavily dependent on an individual's earnings history. The NDPF literature offers a potential rationale for this history dependence as well as an explanation as to why history dependence appears where it does but not more broadly.

One example of a history-dependent transfer program is Social Security Disability Insurance (SSDI). Monthly SSDI benefits are based on pre-disability earnings, such that an individual with a history of higher earnings will receive a larger transfer than an individual with a history of lower earnings. Another example is unemployment insurance (UI). Most state UI systems are designed to replace approximately 50 percent of an individual's pre-unemployment

wages up to a maximum amount (e.g., \$504 per week in New York as of this writing).³⁷ Thus, high-lows receive larger transfers than low-lows when low earnings in period 2 are due to unemployment.

On first glance, the history dependence of SSDI and UI is puzzling. Why should the fact that someone earned more in the past entitle them to a *larger* transfer today? As the political theorist Robert Goodin observes: “We know, from their tax policies, that many governments themselves think that a more equal pattern of income distribution would be preferable and that it is government’s job to promote it. Yet the very same governments, through their compensation policies, set systematically about reproducing the same nonideal pattern of income distribution that they try to correct through their tax policies.”³⁸ This “well-nigh universal” practice, according to Goodin, is “in certain ways flatly contrary to ... public judgments made in other contexts about the justice of income differentials.”³⁹

The NDPF literature suggests an answer to Goodin’s riddle. SSDI and UI benefits affect the implicit tax rate on prior-period labor income. One of the benefits of earning a high wage in period 1 is that it entitles you to a larger transfer if you become disabled or unemployed in period 2. The history dependence of SSDI and UI benefits thus relaxes the incentive compatibility constraint in period 1, potentially allowing the government to impose higher labor income taxes without inducing mimicking. To be sure, these improved period 1 incentives come at the cost of period 2 incentives, since a larger period 2 transfer makes the high-low strategy more attractive for high- θ types. However, the linkage of SSDI and UI benefits to discrete negative θ shocks reduces the risk of period 2 mimicking: while a high- θ type may be able to fake a disability or induce her employer to fire her, this is certainly harder than simply earning a low period 2 wage. Moreover, if high- θ types are risk averse and the government can verify the negative θ shock (disability or involuntary termination) with reasonable accuracy, then the value of the additional insurance to high- θ types who earn high wages in period 1 will exceed the cost to the government of providing the insurance. Larger transfers to high-lows in SSDI and UI can therefore allow for higher tax rates on high wage earners in period 1, benefitting low-lows as well.

Crucial to this justification for history-dependent SSDI and UI benefits is the assumption that the government can verify disability or involuntary termination. Verification—even if imperfect—reduces the period 2 moral hazard for high- θ types of taxes that are regressive over past labor income. Without verification, it is theoretically ambiguous whether the positive effect of past-income regressivity on period 1 incentives outweighs the negative effect on

³⁷ U.S. Dep’t of Labor, Employment & Training Admin., Office of Unemployment Ins., Significant Provisions of State Unemployment Insurance Laws Effective January 2021, <https://oui.doleta.gov/unemploy/content/sigpros/2020-2029/January2021.pdf> (last updated May 10, 2021).

³⁸ Robert E. Goodin, Compensation and Redistribution, 33 *Nomos* 143, 169 n.18 (1991).

³⁹ *Id.* at 147.

period 2 incentives. Thus, the justification for history-dependent SSDI and UI benefits would not necessarily apply to all other proposals for history dependence (e.g., income averaging).⁴⁰

In sum, the NDPF literature suggests a reason why governments might choose to incorporate history dependence—and specifically, past-income regressivity—into programs that provide transfers conditional upon semi-verifiable negative θ shocks. It therefore rationalizes the pattern of history dependence in these programs with history independence in income taxation more broadly. This is not to suggest that current SSDI and UI benefits are optimal in all respects (indeed, they almost certainly are not⁴¹). But it does suggest that an incongruity between SSDI/UI and other elements of the tax-and-transfer system—the regressive history dependence of SSDI/UI versus the progressivity and history independence of general income taxation—may not be so incongruous after all.

Capital Taxation. NDPF’s prescription for capital income tax rates that are regressive over labor income may seem at odds with the federal statutory rate structure, which effectively imposes capital income tax rates that rise with labor income. (Capital income tax rates—i.e., tax rates on interest, dividends, and capital gains—depend on taxable income, which in turn encompasses both capital and labor income.) However, the full tax-and-transfer system includes several examples of higher capital taxes on lower labor-income earners. The NDPF literature provides a potential rationale—at least at a theoretical level—for these otherwise-puzzling elements.

One example of a labor-income-regressive capital income tax is the earned income tax credit’s investment income limit, which denies the credit to individuals with more than \$3,650 of investment income (interest, dividends, capital gains, and certain other sources of nonbusiness income).⁴² The investment income limit functions as an astronomically high marginal tax rate on the 3,651st dollar of capital income for individuals with low labor income (a 672,800 percent rate for individuals claiming the maximum credit).⁴³ A second example is Supplemental Security Income (SSI), which is available only to individuals with “countable resources” of \$2000 or less (\$3000 for a couple). Countable resources include stocks and bonds, bank accounts, motor vehicles (with a one-vehicle exemption), and real estate (with an exemption for one personal residence).⁴⁴ A similar asset test applies to some Medicaid

⁴⁰ See, e.g., Lily L. Batchelder, *Taxing the Poor: Income Averaging Reconsidered*, 40 *Harv. J. on Legislation* 395 (2003) (proposing that individuals whose incomes decline over a two-year period be allowed to carry back the unused portion of their standard deduction and personal exemption from the later year to the earlier year, effectively allowing high-lows to receive a larger transfer in period 2 than low-lows).

⁴¹ On flaws in the design of UI systems, see Brian Galle, *How to Save Unemployment Insurance*, 50 *Ariz. St. L.J.* 1009 (2018).

⁴² See I.R.C. § 32(i).

⁴³ See Internal Revenue Service, *Earned Income and Earned Income Tax Credit (EITC) Tables*, <https://www.irs.gov/credits-deductions/individuals/earned-income-tax-credit/earned-income-and-earned-income-tax-credit-eitc-tables> (last visited Sept. 14, 2021) (maximum credit of \$6,728 in 2021 for taxpayers with three or more qualifying children).

⁴⁴ Social Security Administration, *Spotlight on Resources—2021 Edition*, <https://www.ssa.gov/ssi/spotlights/spot-resources.htm> (last visited Sept. 14, 2021).

beneficiaries.⁴⁵ These asset limits function as wealth (rather than capital income) taxes on individuals with low labor earnings.

The NDPF literature helps to explain why policymakers might choose to include implicit capital income and wealth taxes in transfer programs such as the EITC, SSI, and Medicaid. These implicit capital taxes discourage high- θ types from playing a high-low strategy by making it harder for them to smooth consumption across high-wage and low-wage periods. That, in turn, allows the government to provide larger transfers to low wage earners, who are now likelier to be truly low- θ types. To be sure, the existing EITC investment income limit and the SSI and Medicaid asset limits still may be flawed in various ways (the SSI limit, in particular, is shockingly low). However, the NDPF literature suggests that insofar as these features are flawed, it is not *because* they are regressive over labor income. Capital taxes that are regressive over labor income can play an important role in achieving a more redistributive tax-and-transfer system.

At the same time, the NDPF literature calls into question other features of our capital income tax system. In particular, it raises serious doubts about the structure of traditional retirement accounts (e.g., traditional IRAs and 401(k) plans), which allow taxpayers to claim an immediate deduction for savings and to defer tax until withdrawal. Recall that cash-flow treatment—an immediate deduction for savings with a tax on dissavings—is equivalent to a consumption tax, which is roughly equivalent to the nontaxation of capital income. However, when marginal rates rise (fall) between the time of contribution and the time of withdrawal, cash-flow taxation produces a positive (negative) effective tax rate on savings. Thus, traditional IRAs and 401(k) plans result in a capital tax or a capital subsidy depending on whether an individual’s marginal tax rate rises or falls. An individual whose earnings increase between the time of contribution and the time of withdrawal will generally move into a higher marginal-rate bracket, and thus the traditional IRA or 401(k) plan will result in a positive capital income tax. An individual whose earnings decrease will generally move into a lower bracket, and thus the traditional IRA or 401(k) plan will result in a negative capital income tax.

In this respect, traditional IRAs and 401(k) plans—like the EITC investment income limit and the SSI and Medicaid asset tests—impose capital income taxes that depend on labor income. But for traditional IRAs and 401(k) plans, the capital income tax depends on labor income in precisely the wrong way. The individuals who receive capital subsidies are the high-lows, who are exactly the people on whom we want to impose positive capital *taxes*. The traditional IRA/401(k) structure facilitates consumption smoothing by would-be mimickers, which is precisely what NDPF seeks to prevent.

The NDPF literature thus sheds light on the long-running contest between traditional and Roth retirement accounts. Although the NDPF literature is skeptical of retirement savings

⁴⁵ In New York, for example, the Medicaid asset test applies to individuals who otherwise qualify for Medicaid by virtue of being blind, disabled, or age 65 and over. See N.Y. State Dep’t of Health, Medicaid in New York State, https://www.health.ny.gov/health_care/medicaid (last visited Sept. 14, 2021).

incentives in general (and, indeed, suggests that the government ought to *discourage* savings under some circumstances), NDPF insights provide a strong reason for policymakers to favor Roth plans over traditional plans if forced to make a choice. Traditional IRAs and 401(k) plans subsidize saving by exactly the individuals whose savings we should discourage: individuals who earn high incomes in “period 1” (i.e., before they turn 59 ½ and become eligible for tax-free IRA withdrawals) and earn low incomes in “period 2” (after age 59 ½). Traditional IRAs and 401(k) plans therefore exacerbate moral hazard among older individuals and make it costlier for the government to provide comprehensive θ -shock insurance over the lifecycle.

Age Dependence. Statutory income tax rates do not explicitly depend upon age. However, the existing federal income tax incorporates elements of age dependence. For example, the combination of progressive marginal rates and increasing income over the lifecycle means that average rates will typically rise for taxpayers into middle age, in accordance with NDPF’s prescriptions. Likewise, though of a much smaller magnitude, the additional \$600 standard deduction for individuals 65 and older slightly reduces average rates for senior citizens, in accordance with NDPF’s suggestion that rates should decline when the price elasticity of labor supply sharply rises.

Other age-dependent elements of the tax-and-transfer system align less well with the lessons of NDPF. For example, the EITC for taxpayers without qualifying children is typically limited to individuals ages 25 to 64.⁴⁶ (These age restrictions are temporarily suspended for tax year 2021.⁴⁷) The age restrictions that normally apply are difficult to rationalize. The EITC—though it has cross-cutting effects on marginal rates—always reduces the effective average tax rate on labor income when it applies. The NDPF literature suggests that tax rates on labor income should be lower for early adults and older workers; the EITC age restrictions accomplish the exact opposite.

Age-dependent taxation remains an area ripe for research. Although scholars are beginning to reach consensus on the optimal shape of the age-dependent tax schedule, the implementation challenges of age-dependent taxes are understudied.⁴⁸ We will return to this issue in Part III, when we consider the particular challenges that age-dependent taxation poses for contract law in the employment context.

Value-Added Taxation. Finally, the NDPF literature offers insights for the debate over value-added taxes (VATs), a feature of the tax system in virtually every high-income country other than the United States. VATs are consumption taxes; consumption taxation is roughly equivalent to zero capital income taxation; and the NDPF literature supplies arguments in *favor* of capital income taxation (at least for a subset of the population). One might therefore think that NDPF provides an argument *against* VATs—a suggestion that the United States, almost

⁴⁶ I.R.C. § 32(c)(1)(A)(ii)(II).

⁴⁷ *Id.* § 32(n).

⁴⁸ For an important—and prescient—exception, see Lee Anne Fennell & Kirk J. Stark, *Taxation Over Time*, 59 *Tax Law Review* 1 (2005).

alone in the industrialized world, might be doing it right. However, the implications of NDPF for VATs turn out to be more complicated than that, and NDPF offers one (potentially two) arguments *in favor* of value-added taxation.

First, and most robustly, the NDPF literature suggests that a one-time wealth tax *would* be an efficient form of redistribution *if* the government could commit never to do it again. As is now generally acknowledged, the imposition of a VAT entails a one-time tax on existing wealth plus zero net capital taxation going forward.⁴⁹ For example, when the government imposes a 20 percent VAT (on a tax inclusive base), anyone who previously could have purchased \$100 of after-tax consumption now can purchase only \$80. A large portion of the efficiency benefits of a VAT arise from this one-time wealth-tax feature.⁵⁰

The problem with a one-time capital levy is that individuals won't necessarily believe the government's promise not to do the same thing again. Yet for reasons that are as much sociological as economic, individuals do *not* seem to interpret the imposition of a VAT as a declaration of open season for wealth taxation. (We know this because virtually every other major industrialized nation has imposed a VAT, and their VATs have not triggered discernible capital flight.) To be sure, it is possible that Americans will see VATs as a sneaky strategy to impose a one-time capital levy (which, concededly, they are). But the experience of other countries suggests that VATs offer a rare opportunity to impose a one-time capital levy without bearing the reputational costs that a one-time capital levy otherwise would entail.

Second, a total shift from an individual income tax to a VAT could bolster the credibility of the government's commitment not to exploit information about past labor income. With a VAT, the government typically does not know whose consumption is whose. Replacing the individual income tax with a VAT would potentially disable the government from imposing retrospective labor income taxes because the government would no longer be collecting information on individuals' labor incomes.⁵¹ This feature of a VAT will be particularly attractive to governments that otherwise would not be able to make credible non-exploitation commitments.

In sum, while the NDPF literature challenges the zero capital tax result, it potentially supports one of the policy prescriptions flowing from the zero capital tax result: value-added taxation. While NDPF can justify many puzzling features of the U.S. tax-and-transfer system, the puzzling absence of a VAT in the United States is not one of them.

III. Implications for Other Areas of Law

⁴⁹ See David Altig, Alan J. Auerbach, Laurence J. Kotlikoff, Kent A. Smetters & Jan Walliser, *Simulating Fundamental Tax Reform in the United States*, 91 *Am. Econ. Rev.* 574, 575-76 (2001).

⁵⁰ *Id.*

⁵¹ Note that a flat rate VAT could be coupled with a system of demogrants in order to achieve average-income progressivity—still without the government knowing individuals' labor incomes.

The law-and-NDPF project primarily seeks to shed light on the tax-and-transfer system, but its ambitions and implications extend to other realms as well. This final part draws lessons from NDPF for non-tax areas of law—focusing on contracts, torts, property, and constitutional law.

Contracts. NDPF’s prescriptions for history-dependent and age-dependent taxation implicitly assume that an individual’s wage is equal to the marginal product of her labor. Long-term employment contracts can shake that assumption. For example, if tax rates are hump-shaped over the age distribution, then younger workers will have an incentive to accelerate wages and older workers will have an incentive to defer wages. Insofar as income-shifting techniques succeed, they pose a serious threat to history-dependent and age-dependent taxes.

Contract law can facilitate or frustrate income-shifting. One way that the common law of contracts frustrates income shifting is through its refusal to order specific performance of employment agreements.⁵² Imagine, for example, that a 30-year-old and her employer agree that she will work for the employer for 10 years but receive most of her compensation in years 1, 2, and 3 (when her age-dependent tax rate remains relatively low). If the employee quits in year 4, the employer will have limited recourse. In some states (though not in California), the employer may be able to enforce a noncompete agreement that temporarily prevents the employee from working elsewhere. In no state, though, will a court order the employee to go back to work. By withholding specific performance in the employment context, the common law generates frictions that make it more difficult for individuals to manipulate history-dependent and age-dependent taxes. In this instance, the common law is complementary to NDPF-inspired taxation.

Now imagine that a 55-year-old and her employer agree that she will work for the employer for 10 more years and receive most of her compensation in year 10 (when her age-dependent tax rate has declined again). Under these circumstances, the employer’s commitment to pay a large year-10 wage likely would be enforceable in court (since the remedy would be a money judgment rather than involuntary servitude). The enforceability of long-term employment contracts *against the employer* creates a tension between the common law of contracts and the new dynamic public finance. Contract law, at least as it exists today, is more conducive to tax rates that rise over time than to tax rates that fall.

Torts. One of the most important questions in tort law is whether and how the victim’s income ought to affect damages. Income dependence in the context of tort damages raises similar issues as history dependence in the context of SSDI and UI. The opportunity to collect larger damages in the event of an injury is an additional benefit of earning a high period 1 wage. The additional benefit eases the incentive compatibility constraint with respect to high- θ types in period 1, thus allowing the government to further increase the period 1 tax rate on high wage earners.

⁵² See *Lumley v Wagner* [1852] EWHC (Ch) J96 (noting that court cannot compel mezzo-soprano singer to perform at Her Majesty’s Theatre).

In tort law, as in the SSDI and UI cases, regressivity with respect to past income is accompanied by a verification mechanism. The verification mechanism in the tort law context is the jury. When the jury finds for the victim, it indicates that the victim has indeed experienced a negative θ shock. This verification step, in turn, increases the likelihood that the individuals who benefit from regressivity with respect to past income in the tort context are genuine high-lows rather than mimickers.

Property. One of property law's central concerns is the legibility of holdings.⁵³ "Legibility" in this context refers to susceptibility to observation and taxation. As the political scientist and anthropologist James C. Scott argues, the desire for legibility played a significant role in the rise of the freehold tenure system across early modern Europe.⁵⁴ Today, efforts inspired by the Peruvian economist Hernando de Soto to give formal legal title to property holders in "slums" seek in large part to enhance legibility.

The new dynamic public finance underscores legibility's double-edged quality. Legibility aids capital taxation, and the NDPF literature offers justifications for certain forms of capital taxation. NDPF thus emphasizes the importance of legibility. On the other hand, legibility can facilitate capital expropriation, and the NDPF literature highlights how hard it can be for governments to make credible non-expropriation commitments. NDPF thus reveals a potential cost of legibility: a shadow tax on capital as long as an expropriation risk remains.

Illegibility is a property-law analogue to anonymity in financial markets. When holdings are illegible, as they were throughout much of medieval Europe, the government does not know whose is whose. Whether a government must resort to illegibility/anonymity in order to reduce shadow taxes on capital depends in large part on the strength of its other institutions. In this respect, property institutions and constitutional checks on expropriation are complements.

Constitutional Law. Finally, the NDPF literature offers insights for ongoing debates in U.S. constitutional law. One of those debates concerns wealth taxation. Article I of the Constitution requires that any "direct tax" be apportioned among the states based on population—a requirement that, if applied to wealth taxation, would result in huge state-to-state variations in tax rates.⁵⁵ Scholars who argue that wealth taxes shouldn't be classified as

⁵³ See David Weisbach & Daniel Hemel, *The Legal Envelope Theorem*, *Boston University Law Review* (forthcoming 2022), <https://ssrn.com/abstract=3725155>.

⁵⁴ James C. Scott, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* 40 (Yale Univ. Press, 1998).

⁵⁵ For example, if the tax were applied to households with wealth of \$5 million or more, the tax rate in West Virginia would be approximately 16 times the rate in Connecticut. See Daniel Hemel, *How To Tax Wealth Constitutionally*, *Medium: Whatever Source Derived* (Jan. 28, 2019), <https://medium.com/whatever-source-derived/how-to-tax-wealth-constitutionally-863ce992ac7e>; Alex Zhang, *The Wealth Tax: Apportionment, Federalism, and Constitutionality*, 23 *U. Pa. J. Law & Soc. Change* 269, 283 tbl.1 (2020). [Add Brooks & Gamage]

“direct taxes” emphasize the importance of giving Congress “flexibility in meeting twenty-first century challenges.”⁵⁶

The NDPF literature, while offering arguments in support of capital taxation, raises some doubts about the benefits of “flexibility.” One way that a government might credibly commit not to impose a large capital levy is through constitutional *inflexibility*. A constitutional rule that freely allows Congress to impose capital income taxes but severely restricts wealth taxes would make some sense if we think that capital taxation is desirable in moderation but fear of limitless capital taxation may lead to large distortions. A capital income tax—when implemented on an accrual basis—functions more or less like a wealth tax with a rate capped by the accrual rate.⁵⁷ Maintaining the implicit rate cap is arguably desirable as a commitment device.

The NDPF literature also bears implications for a related debate over retrospective labor income taxation. In a trio of late 1920s decisions—*Nichols v. Coolidge*,⁵⁸ *Blodgett v. Holden*,⁵⁹ and *Untermeyer v. Anderson*⁶⁰—the Supreme Court held that certain estate and gift tax provisions could not be applied to pre-enactment transfers. The precedential weight of the cases is now quite uncertain, but the Court has never overruled them explicitly.⁶¹

Nichols, *Blodgett*, and *Untermeyer* may be artifacts of the late *Lochner* era, but they respond to an ongoing concern: the possibility that Congress will impose new taxes based on past-period behavior (and especially for our purposes, past-period labor income). One takeaway from the NDPF literature is that the problem of retrospective labor income taxation shares the same essential structure as the more-discussed problem of capital expropriation. The upshot is not necessarily that courts should tie Congress’s hands—the Constitution does not enact the New Dynamic Public Finance any more than it enacts Mr. Herbert Spencer’s Social Statics.⁶² However, NDPF helps us understand why flexibility might not always be a virtue.

Conclusion

Like other joiners of law and economics, law and the new dynamic public finance does not yield a single set of unambiguous policy prescriptions. But it does reveal unappreciated justifications for existing policies and calls attention to cases in which current law may be misfiring. The conclusions don’t fit neatly within existing ideological categories; they are not always comfortable; and they may not all turn out—after further inquiry—to be right. But a

⁵⁶ Dawn Johnsen & Walter Dellinger, *The Constitutionality of a National Wealth Tax*, 93 Ind. L.J. 111, 115 (2018).

⁵⁷ Daniel Hemel, *Taxing Wealth in an Uncertain World*, 72 Nat’l Tax J. 755 (2019).

⁵⁸ 274 U.S. 531 (1927).

⁵⁹ 275 U.S. 142 (1927).

⁶⁰ 276 U.S. 440 (1928).

⁶¹ See *United States v. Carlton*, 512 U.S. 26, 34 (1994) (stating that *Nichols*, *Blodgett*, and *Untermeyer* “were decided during an era characterized by exacting review of economic legislation under an approach that ‘has long since been discarded,’” and adding that “[t]o the extent that their authority survives, they do not control here”).

⁶² Cf. *Lochner v. New York*, 198 U.S. 45, 75 (1905) (Holmes, J., dissenting) (“The Fourteenth Amendment does not enact Mr. Herbert Spencer’s Social Statics.”).

serious effort to address the problem of income inequality over the lifecycle will likely require sustained engagement with NDPF's insights. And at the very least, even if NDPF does not lead the way to a more egalitarian future, it promises a richer understanding of the tax-and-transfer system that we already have.