Law and the New Dynamic Public Finance

Daniel J. Hemel†

In recent years, economists working in public finance and related fields have applied a new set of models and methodologies to the problems of inequality and economic insecurity. This innovative approach—often known as "the new dynamic public finance"—emphasizes the challenges posed by productivity changes across the life cycle and policy changes over time. The new dynamic public finance has generated important insights that often cut against the conventional wisdom in classical optimal tax theory, suggesting—for example—that capital income should be taxed and that higher capital income taxes can incentivize investment under certain circumstances. Mainstream economics has assimilated many of these insights, but with fleetingly few exceptions, legal scholars have yet to engage with the new dynamic public finance literature.

This article explores the potential for cross-pollination between law and the new dynamic public finance. It explains the central intuitions underlying dynamic tax models and draws out implications for tax and other areas of law. The new dynamic public finance offers compelling reasons to adopt age-dependent tax schedules, generates novel justifications for the taxation of capital, and provides an original argument for applying different tax rates to single individuals and married couples. It also reveals flaws in the federal taxation of retirement savings and—in particular—raises serious doubts about the wisdom of the traditional IRA and 401(k) structures. Beyond the Internal Revenue Code, the new dynamic public finance literature offers fresh perspectives on the design of disability and unemployment insurance programs, the rationale for tort law, the regulation of cryptocurrency, and the application of the Constitution's procedural due process requirements, among other subjects. Ultimately, incorporating a new dynamic public finance perspective into tax and non-tax law can inform the crafting of a more comprehensive system of social insurance while enriching our understanding of the system we now have.

---

† John S. R. Shad Professor of Law, New York University School of Law. 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 Stefan Bechtold, Dhammika Dharmapala, Michael Doran, David Elkins, Joshua Fischman, William Gale, Peter Ganong, Michael Gilbert, Jacob Goldin, Andrew Hayashi, Deborah Hellman, Rich Hynes, Paul Mahoney, Ruth Mason, Noah Marks, David Mitchell, Caleb Nelson, Richard Schmalbeck, Daniel Shaviro, Steven Sheffrin, Megan Stevenson, Ethan Yale, Lawrence Zelenak, participants in workshops at Duke Law School, ETH Zürich, New York University, the University of Chicago Law School, and the University of Virginia School of Law, and attendees at the National Tax Association 118th Annual Conference on Taxation in Boston. All errors are my own.

Introduction

Economic inequality is—in former President Obama's words—"the defining challenge of our time."[1] Economic inequality is also, in important respects, a problem *of* time. Within any given age cohort, economic inequality grows over time: the distribution of income is highly egalitarian among individuals in their twenties and becomes increasingly lopsided as individuals move into middle and older age.[2] Across generations, economic inequality can compound over time, rigidifying class distinctions and potentially threatening social stability.[3] Meanwhile, efforts to address economic inequality through the tax-and-transfer system are complicated by the inconsistency of government policies over time—the inability of lawmakers in one period to commit to particular policies in the next. For all these reasons, economists have come to understand and analyze inequality and taxation as fundamentally dynamic phenomena—phenomena characterized by change over time.

Notwithstanding the centrality of time to inequality and taxation, the two most influential models of optimal taxation in the U.S. tax law

1. Press Release, White House Office of the Press Sec'y, Remarks by the President on Economic Mobility (Dec. 4, 2013, 11:31 AM), https://obamawhitehouse.archives.gov/the-press-office/2013/12/04/remarks-president-economic-mobility [https://perma.cc/T6TJ-EDVT].

2. *See infra* Figure 1 and text accompanying nn. 75-76.

3. *See* THOMAS PIKETTY, CAPITAL IN THE TWENTY-FIRST CENTURY 1 (Arthur Goldhammer trans., 2014).

literature—the static Mirrlees model[4] and the Atkinson-Stiglitz model[5]—devote scant attention to problems of time. The static Mirrlees model—which dates back to a 1971 article by the British economist James Mirrlees—is, in its author's own words, "timeless."[6] Although the article was pathbreaking in other respects, Mirrlees acknowledged that "[i]ntertemporal problems are ignored" in his analysis.[7] The 1976 Atkinson-Stiglitz model, for its part, flattens the dimension of time. Its key analytical move is to transform consumption at different times into consumption of different goods at the same time.[8] The complexities of time—in particular, changes over time in individual skill levels and in government policies—are assumed away.

Optimal tax theory has made considerable progress since 1976. Under the banner of "the new dynamic public finance" (NDPF),[9] these scholars—including Mirrlees himself before his death in 2018—have incorporated a much richer conception of change over time into the earlier workhorse models of optimal taxation, generating important insights into the problems of economic inequality and insecurity. In some cases, this new approach has upended the conventional wisdom in optimal tax theory. For example, the main capital-tax-related takeaway from the Atkinson-Stiglitz model—that capital should face a zero rate of tax—has not survived NDPF analysis. Beyond that insight, NDPF models have yielded several other policy prescriptions with real-world relevance. For example:

- Capital taxes should depend on labor income and earnings history, with "high-lows"—individuals who report high wages in earlier periods and low wages in later periods—facing higher capital tax rates;[10]
- Labor income tax schedules should be age-dependent, with progressivity increasing from early adulthood through midlife but falling around retirement age;[11]

4. *See* J.A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38 REV. ECON. STUD. 175, 176-78 (1971).

5. A.B. Atkinson & J.E. Stiglitz, *The Design of Tax Structure: Direct versus Indirect Taxation*, 6 J. PUB. ECON. 55, 67-70 (1976) (modeling consumption at different dates as equivalent to consumption of different goods at a single point in time, thereby abstracting from temporal dynamics).

6. *See* Mirrlees, *supra* note 4, at 175.

7. *See id.*

8. *See* Atkinson & Stiglitz, *supra* note 5, at 69.

9. 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 that suggested by my own individual papers," Kocherlakota writes. NARAYANA R. KOCHERLAKOTA, THE NEW DYNAMIC PUBLIC FINANCE 2 n.1 (Princeton University Press, 2010). The title he chose was "deliberately catchy." *Id.*

10. *See infra* Section II.A.

11. *See infra* Section II.C.

- Single individuals and married couples should face different tax schedules, with less progressive tax schedules for married couples than for singles;[12] and
- At least under some circumstances, capital taxation—rather than entailing a tradeoff between efficiency and equity—can enhance efficiency while advancing equity by bolstering the credibility of government policy.[13]

Mainstream economics has already assimilated key conclusions from the new dynamic public finance. All the top peer-reviewed economics journals have published NDPF papers.[14] Undergraduate-level economics texts cover basic NDPF concepts.[15] Yet legal scholarship has largely proceeded as if the dynamic turn in public economics never happened. In a 2007 *Stanford Law Review* article, Daniel Shaviro presciently noted that the NDPF literature, though "as yet little known to legal scholars," contained "important implications" for tax law scholarship.[16] Since 2007, though, only two articles in Westlaw's database of law reviews have even mentioned "the new dynamic public finance" in body text,[17] and none has sought to take stock of NDPF's wide-ranging implications for legal analysis.[18]

This article attempts to bridge the divide between law and the new dynamic public finance, providing a nontechnical introduction to key

12. *See infra* Section II.D.

13. *See infra* text accompanying notes 91-92.

14. *See, e.g.*, Mikhail Golosov, Narayana Kocherlakota & Aleh Tsyvinski, *Optimal Indirect and Capital Taxation*, 70 REV. ECON. STUD. 569, 569-87 (2003); Mikhail Golosov, Maxim Troshkin & Aleh Tsyvinski, *Redistribution and Social Insurance*, 106 AM. ECON. REV. 359, 359-96 (2016); Mikhail Golosov & Aleh Tsyvinski, *Designing Optimal Disability Insurance: A Case for Asset Testing*, 114 J. POL. ECON. 257, 257-303 (2006); Narayana R. Kocherlakota, *Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation*, 73 ECONOMETRICA 1587, 1587-1614 (2005); Iván Werning, *Optimal Fiscal Policy with Redistribution*, 122 Q.J. ECON. 925, 925-68 (2007). For recent reviews of the NDPF literature, see Mikhail Golosov & Aleh Tsyvinski, *Policy Implications of Dynamic Public Finance*, 7 ANN. REV. ECON. 147, 147-173 (2015); and Stefanie Stantcheva, *Dynamic Taxation*, 12 ANN. REV. ECON. 801, 801-36 (2020).

15. *See, e.g.*, KARTIK B. ATHREYA, BIG IDEAS IN MACROECONOMICS: A NONTECHNICAL VIEW 154-57 (2013); JONATHAN GRUBER, PUBLIC FINANCE AND PUBLIC POLICY 73-77 (7th ed. 2022); BERNARD SALANIÉ, THE ECONOMICS OF TAXATION 134-36 (2d ed. 2011).

16. *See* Daniel Shaviro, *Beyond the Pro-Consumption Tax Consensus*, 60 STAN. L. REV. 745, 751-58 (2007).

17. *See* Joseph Bankman & Daniel Shaviro, *Piketty in America: A Tale of Two Literatures*, 68 TAX L. REV. 453, 464 (2015); Edward D. Kleinbard, *Capital Taxation in an Age of Inequality*, 90 S. CAL. L. REV. 593, 637 (2017).

18. Legal scholar and economist Louis Kaplow incorporates NDPF insights at several points in his 2008 volume on public finance, *see* LOUIS KAPLOW, THE THEORY OF TAXATION AND PUBLIC ECONOMICS 160, 227-29, 241-43 (2008), but Kaplow's book is primarily directed at scholars and students of economics rather than law, *see id.* at 160. Legal scholars Lee Anne Fennell and Kirk Stark discuss age-based taxation in an insightful 2005 article. *See generally* Lee Anne Fennell & Kirk J. Stark, *Taxation Over Time*, 59 TAX L. REV. 1 (2005) (disucssing this theme). They even anticipate some of the arguments that the NDPF literature would later develop. *See, e.g.*, *id.* at 38-40 (noting potential efficiency benefits of tax rates that increase from young adulthood into middle age). However, at the time that Fennell and Stark were writing, the NDPF literature was still in its nascent stage, and their article does not mention or discuss the new dynamic public finance.

NDPF concepts and highlighting the most important implications of NDPF for law and policy. While insights from the NDPF literature are especially relevant to tax law, NDPF also yields lessons for other areas, including Social Security Disability Insurance (SSDI) and unemployment insurance (UI), torts, contracts, property, financial regulation, and constitutional law. In some cases, NDPF analysis supplies possible justifications for existing (though puzzling) features of current law in the United States and elsewhere, such as the system of history-dependent labor income taxes implicit in disability insurance and unemployment insurance programs. In other cases, as with the regulation of cryptocurrency, NDPF analysis clarifies tradeoffs without yielding definitive answers. And beyond its implications for specific areas of law, NDPF sheds light on one of the central debates in late twentieth and early twenty-first century law and economics: whether non-tax tools should be used for redistributive ends.[19] The NDPF literature highlights a variety of ways in which areas of law other than tax can contribute to the tax system's redistributive and social-insurance missions.

Part I of this article provides an overview of the NDPF approach. Part II explores the implications of NDPF analysis for tax-and-transfer policy. Part III considers connections to other areas of law. A conclusion highlights promising paths for future work.

I. An Introduction to the New Dynamic Public Finance

A. First-Generation Optimal Tax Theory

To explain the contributions of the new dynamic public finance, it will be helpful to start by briefly surveying the landscape of optimal tax theory as it existed prior to the arrival of NDPF. Two paradigms—the "static"

19. *See generally* Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL STUD. 667 (1994) (arguing that redistributive goals are better pursued through the tax system than through legal rules); Christine Jolls, *Behavioral Economics Analysis of Redistributive Legal Rules*, 51 VAND. L. REV. 1653 (1998) (applying behavioral-economics insights to assess whether legal rules can complement redistribution through taxation); Chris William Sanchirico, *Taxes Versus Legal Rules as Instruments for Equity: A More Equitable View*, 29 J. LEGAL STUD. 797 (2000) (reconsidering the Kaplow-Shavell framework and arguing for a more nuanced role for legal rules in redistribution); Louis Kaplow & Steven Shavell, *Should Legal Rules Favor the Poor? Clarifying the Role of Legal Rules and the Income Tax in Redistributing Income*, 29 J. LEGAL STUD. 821 (2000) (clarifying that redistribution should generally be accomplished through the income-tax system, not through substantive legal doctrines); David A. Weisbach, *Should Legal Rules Be Used to Redistribute Income?*, 70 U. CHI. L. REV. 439 (2003) (arguing that attempts to redistribute through legal rules are inefficient and duplicative of tax-transfer mechanisms); Ronen Avraham, David Fortus & Kyle Logue, *Revisiting the Role of Legal Rules and Tax Rules in Income Redistribution: A Response to Kaplow & Shavell*, 89 IOWA L. REV. 1125, 1130 (2004) (contending that under certain conditions, legal rules can supplement tax-based redistribution); Lee Anne Fennell & Richard H. McAdams, *The Distributive Deficit in Law and Economics*, 100 MINN. L. REV. 1051 (2016) (critiquing efficiency-focused law-and-economics analysis for neglecting distributional effects); Zachary Liscow, *Is Efficiency Biased?*, 85 U. CHI. L. REV. 1649 (2018) (arguing that conventional efficiency analysis systematically favors the wealthy and should be supplemented with distributive considerations).

Mirrleesian model of labor income taxation and the Atkinson-Stiglitz model of capital taxation—have exerted enormous influence over economics and economically informed legal scholarship over the last half century. Mirrlees's 1971 article has garnered more than 7,500 citations in economics journals and more than 170 citations in the Westlaw database of law-related secondary sources[20]—an unusually high number for an economics paper—and formed part of the basis for Mirrlees's 1996 Nobel Prize.[21] Atkinson and Stiglitz's 1976 article has generated more than 2,800 citations in economics journals and roughly seventy-five citations in legal secondary sources[22]—and is also connected to the work for which Stiglitz won a 2001 Nobel Prize.[23] Like the static Mirrleesian model, the Atkinson-Stiglitz theorem is "considered to be one of the most important advances of the last century of public-finance economics."[24]

1. Mirrlees and Rawls

Optimal tax theory's main focus—dating back to Mirrlees's 1971 article[25]—is the construction of a welfare-maximizing tax schedule when the government can observe individuals' income but not their labor effort or "skill." A brief overview of the 1971 model will aid the analysis that follows.

Mirrlees starts from the familiar premise of diminishing marginal utility of consumption: an extra unit of consumption brings less utility (or happiness) as one climbs the income distribution.[26] The government in Mirrlees's model is benevolent and seeks to maximize social welfare, with welfare defined as a function of individual utilities.[27] The government therefore wants to redistribute from high-income individuals (for whom the marginal utility of consumption is low) to low-income individuals (for whom the marginal utility of consumption is high). Moving a dollar (or for

20. Citation counts are based on Google Scholar and Westlaw searches as of November 2025. For Mirrlees's article, the Westlaw count (170) includes citations to "An Exploration in the Theory of Optim*um* Income Taxation" (the correct title) and citations to "An Exploration in the Theory of Optim*al* Income Taxation" (to which seventeen law-review articles mistakenly cite).

21. *See* Agnar Sandmo, *Asymmetric Information and Public Economics: The Mirrlees-Vickrey Nobel Prize*, 13 J. ECON. PERSPS. 165, 167-169 (1999).

22. *See supra* note 20.

23. *See* Joseph E. Stiglitz, *Information and the Change in the Paradigm in Economics*, 92 AM. ECON. REV. 460, 483-84 & n.27 (2002).

24. David Gamage, *How Should Governments Promote Distributive Justice?: A Framework for Analyzing the Optimal Choice of Tax Instruments*, 68 TAX L. REV. 1, 34 (2014).

25. *See* Mirrlees, *supra* note 4, at 176-78.

26. *Id.* at 176. Because Mirrlees's model is timeless, no income is carried over from a previous period or saved for a future one. Thus, income and consumption are equivalent.

27. *See id.* at 178. A utilitarian social-welfare function—in which welfare is the sum of individual utilities—is one possible welfare function in Mirrlees's model, but the model can accommodate other welfare functions as well (for example, prioritarian social-welfare functions that accord greater weight to the welfare of the worst-off members of society). *See, e.g.*, Matti Tuomala & Matthew Weinzierl, *Prioritarianism and Optimal Taxation*, *in* PRIORITARIANISM IN PRACTICE 72, 73-75 (Matthew Adler & Ole Norheim eds., 2022).

the Scottish-born and Oxford-based Mirrlees, a pound) from a high-income person to a low-income person increases overall welfare.

If labor supply were entirely inelastic—in other words, if people worked the same amount regardless of taxes—then the government's problem would be straightforward: the optimal tax system would entail a 100% income tax rate and a totally egalitarian distribution of resources. The problem facing the government is that when it redistributes more from high-income individuals to low-income individuals, high-skill individuals have less of an incentive to earn and report high incomes. The elasticity of labor supply, combined with the government's inability to observe skill directly, animates the challenge of optimal taxation.

Key to the static Mirrleesian model is the notion of "skill," represented by Mirrlees as "n"[28] and in later work by the Greek letter theta (θ). Mathematically, in Mirrlees's model, n (or θ) is simply the ratio of a person's pre-tax income over the number of hours she works (in other words, her wage).[29] This is a convenient simplification, but as Agnar Sandmo notes, one must interpret θ "more broadly" in order to make sense of the static Mirrleesian model.[30] Some jobs require different amounts of effort per hour: for example, an hour on the job as a longshore worker unloading cargo containers is not necessarily equal to an hour on the job as a law-firm associate reviewing documents. Moreover, some jobs require a significant amount of uncompensated time devoted to skill acquisition (e.g., three years of law school). We can usefully think of θ as a measure of the ease with which an individual can transform labor effort into pre-tax income, or even more generally, as a proxy for the amount of labor income that an individual would have earned in a hypothetical world without distortionary taxes. This latter characterization of θ departs from the simple arithmetic of Mirrlees's model, but it resolves some of the complications that come with thinking of θ as a measure of inherent "skill."[31]

---

28. Mirrlees, *supra* note 4, at 176-77.

29. *See id*. at 176.

30. *See* Sandmo, *supra* note 21, at 168.

31. Most significantly, thinking of θ as a proxy for an individual's income in a world without distortionary taxes largely or entirely addresses the Rawlsian objection to optimal tax theory. *See* JOHN RAWLS, JUSTICE AS FAIRNESS: A RESTATEMENT 158 (Erin Kelly ed., 2001) (stating that a tax on θ—or on "native endowments"—"would violate the priority of liberty" because it "would force the more able into those occupations in which earnings were high enough for them to pay off the tax," and thus "would interfere with their liberty to conduct their life within the scope of the principles of justice"). If θ is simply a proxy for the amount of income an individual would have earned in a world without distortionary taxes, an individual who would have been a beachcomber in a world without distortionary taxes would be classified as having a low θ, even if she has the talent to succeed as a Wall Street lawyer. The beachcomber would not, contra Rawls, be forced to work as a Wall Street lawyer in order to satisfy her high θ-tax liability. For overviews of the θ-taxation debate, see Daniel N. Shaviro, *Inequality, Wealth, and Endowment*, 53 TAX L. REV. 397, 414 (2000); Kirk J. Stark, *Enslaving the Beachcomber: Some Thoughts on the Liberty Objection to Endowment Taxation*, 18 CAN. J.L. & JURIS. 47, 49 (2005); and Lawrence Zelenak, *Taxing Endowment*, 55 DUKE L.J. 1145, 1165-69 (2006).

Mirrlees's 1971 article considers several different distributions of θ, but for expositional purposes, it will be convenient to imagine θ as binary (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.[32] 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. The lower left box in Table 1 is blacked out because low-θ individuals lack the option of choosing a high-wage job.

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

	High Wage	**Low Wage**
High-θ	Consumption: $\text{Wage}_{\text{HIGH}}$ – Tax Leisure: Low	Consumption: Wage_{LOW} + Transfer Leisure: High
Low-θ		Consumption: Wage_{LOW} + Transfer Leisure: Low

For present purposes, we will assume that the social-welfare function accords approximately zero weight to additional consumption by high-wage individuals. Such a welfare function might be justified on utilitarian grounds: the highest earners in society—the Elon Musks and Jeff Bezoses—derive essentially no additional happiness from an extra unit of consumption. Or it might be based on the prioritarian intuition that the utility of the worst-off members of society deserves the greatest weight in the welfare calculus. Either way, 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 (in other words, 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 optimal tax rate on high-wage earners is thus bounded by an *incentive-compatibility constraint*: any further tax increase 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 no longer be incentivized to reveal their type.

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

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: philosopher John Rawls's *A Theory of Justice*.[33] Rawls asks us to imagine ourselves in the "original position," behind a "veil of ignorance" that prevents us from knowing the results of the "natural lottery" that will determine the circumstances into which we will be born.[34] If we could purchase θ-shock insurance before we knew the outcome of the initial θ distribution (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 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.[35]

The static Mirrleesian model emphasizes two parameters that go into determining the optimal progressivity of the income tax system. The first is the variance of the distribution of θ values in society.[36] A higher variance (a more unequal distribution of θ values) corresponds to an optimally more progressive tax schedule. If everyone had the same θ value, there would be little need for an income tax at all; the government could raise revenue through lump-sum taxes, or head taxes, that do not depend upon income and do not distort labor choices. Another way to understand the variance intuition is to think of the problem from behind the veil of ignorance: the riskier the natural lottery—the greater the gap between a "good" and "bad" outcome—the more θ-shock insurance we would want to buy. Progressivity is θ-shock insurance.

A second insight from the static Mirrleesian model is that a higher price elasticity of labor supply corresponds to an optimally less progressive tax schedule.[37] In other words, the more responsive that labor-leisure choices are to the tax rate, the less we want to tax. If labor supply were perfectly elastic (in other words, if labor supply dropped to zero in response to a small increase in taxes), then income taxation would be completely ineffective at redistributing wealth, and we wouldn't want any of it. We can think of the price elasticity of labor supply as a measure of the moral hazard of θ-shock insurance.[38]

---

33. JOHN RAWLS, A THEORY OF JUSTICE (Harvard Univ. Press, 1971).

34. *Id*. at 136-42.

35. *Id*. at 1525-33. Note that even if we are less risk-averse behind the veil of ignorance than Rawls imagines, we still might choose to tax the rich at the revenue-maximizing rate simply because of the diminishing marginal utility of consumption.

36. Mirrlees, *supra* note 4, at 207.

37. *See id*. at 202-04.

38. Two other implications of the static Mirrleesian model have generated considerable—and probably excessive—attention: first, that the marginal income-tax rate should be zero at the top of the income distribution, and second, that the marginal income-tax rate should be zero at the bottom of the income distribution. *See e.g.*, Edmund S. Phelps, *Taxation of Wage Income for*

The dispersion of the θ distribution and the price elasticity of labor supply are empirical parameters that vary across societies and across time. As a result, the optimal income tax schedule (according to Mirrlees's logic) also varies across societies and across time. The main takeaway from the Mirrlees model is not that rates should always be set at particular levels; it is that rates should always reflect these two parameters (the variance of the θ distribution and the price elasticity of labor supply). As we shall see, Mirrlees's insight will play an important role in the emergence and evolution of NDPF.

2. The Atkinson-Stiglitz Theorem

Atkinson and Stiglitz's 1976 article incorporates a time dimension into the Mirrleesian model, though the time dimension remains highly stylized.[39] In the Atkinson-Stiglitz model, individuals work in period one and retire in period two. Individuals can save some of their period-one earnings for period-two consumption, but period-one savings can't affect period-two labor supply (because no one works in period two). Atkinson and Stiglitz ask: if we already have an optimal tax on labor income, is it optimal to impose any additional tax on capital (e.g., a wealth tax or a tax on

Economic Justice, 87 Q.J. ECON. 331, 344 (1973) (arguing for a zero rate at the top); Efraim Sadka, *On Income Distribution, Incentive Effects and Optimal Income Taxation*, 43 REV. ECON. STUD. 261, 266 (1976) (advancing a similar claim); Jesús Seade, *On the Shape of Optimal Tax Schedules*, 7 J. PUB. ECON. 204, 213 (1977) (deriving a zero rate at the top and bottom). Mirrlees himself dismissed these zero-rate results as having "little practical value." J.A. Mirrlees, *Optimal Tax Theory: A Synthesis*, 6 J. PUB. ECON. 327, 340 (1976). To understand the zero-rate-at-the-top result, imagine that when the top marginal income tax rate is 70%, the highest earner in society works 2,000 hours per year for $1,000 dollars per hour. At that point, she is indifferent between earning an additional $1,000 per hour pre-tax ($300 after tax) and devoting an additional hour to leisure. Now imagine that the government allows the highest earner to work a 2,001st hour and earn another $1,000 at a 69.9% marginal tax rate. The highest earner, who previously was indifferent between earning an additional $300 after tax and devoting an additional hour to leisure, now chooses work for an additional $301 after tax. She is no worse off, and the government collects an additional $699 in revenue that it otherwise wouldn't have collected. The government keeps on cutting the marginal rate on the highest earner's next increment of income in order to induce her to work an extra hour until finally the marginal rate on the last increment is approximately zero. The zero-rate-at-the-top result is elegant, but it applies only under very restrictive conditions. Specifically, it requires the government to know the level of top earnings in advance. And it only applies to the very top earner, not to the second highest earner. *See* Peter Diamond & Emmanuel Saez, *The Case for a Progressive Tax: From Basic Research to Policy Recommendations*, 25 J. ECON. PERSPS. 165, 173 & nn.11-12 (2011). It is perhaps better understood as the result of a fun thought experiment rather than an actual policy prescription. The zero-rate-at-the-bottom result depends on the assumption that the economy has no zero-income households. *See* Seade, *supra*, at 205. In that case, a positive marginal tax rate on the first dollar of income has a disincentive effect with no corresponding redistributive gain (because everyone must pay the same tax on their first dollar). Again, practical relevance is limited, as there certainly *are* zero-income households in the real world.

39. Atkinson & Stiglitz, *supra* note 5. For a fuller explanation of the Atkinson-Stiglitz model that is accessible to readers with no economics training, see generally Joseph Bankman & David A. Weisbach, *The Superiority of an Ideal Consumption Tax over an Ideal Income Tax*, 58 STAN. L. REV. 1413 (2006).

income generated by saving)? Atkinson and Stiglitz's answer, in a nutshell, is no: the government should allow for tax-free saving.

The Atkinson-Stiglitz result relies on an analogy between a capital tax and a nonuniform commodity tax. Their approach characterizes period-one consumption and period-two consumption as different goods (or commodities) that can both be purchased with period-one labor income. If the interest rate is 10% and the capital income tax rate is 0%, then an individual can substitute $1 of period-one consumption for $1.10 of period-two consumption. Individuals will allocate their period-one labor income between period-one and period-two consumption until, at the margin, they are indifferent between an extra $1 of consumption in period one and an extra $1.10 of consumption in period two.

Within this framework, capital taxation—whether in the form of a capital income tax or a wealth tax—produces two effects. First, it distorts individuals' choices between period-one and period-two consumption. For example, a 20% capital income tax means that $1 of period-one consumption can be substituted for $1.08 of period-two consumption (rather than $1.10). The change in the intertemporal rate of substitution (the rate at which period-one consumption can be substituted for period-two consumption) will affect the consumption bundle that individuals choose. Conditional on an assumption discussed in the margin,[40] the capital tax will cause individuals to allocate a larger portion of their period-one labor income to period-one consumption (and less to period-two consumption) than they would have chosen to allocate in a tax-free world.

Second, and more subtly, capital taxation distorts individuals' choices between period-one leisure and period-one labor. One reason to choose labor over leisure in period one is to have money to spend on consumption in period two. By reducing the amount of period-two consumption that an individual can finance out of period-one labor income, a capital tax makes period-one labor less attractive than it was before.

40. The statement in text assumes that the elasticity of intertemporal substitution (EIS) is greater than one—in other words, the substitution effect dominates the income effect. Martin Holm and coauthors note that "[s]everal influential contributions in macroeconomics and finance assume that the EIS exceeds one," and they estimate the EIS to be around 1.6 based on a quasi-natural experiment involving a Norwegian dividend tax increase. *See* Martin B. Holm, Rustam Jamilov, Marek Jasinski & Plamen Nenov, *Estimating the Elasticity of Intertemporal Substitution Using Dividend Tax News Shocks* 2-3 (Univ. of Oxford, Working Paper, July 2025), https://users.ox.ac.uk/~econ0628/EIS.pdf [https://perma.cc/26P5-BQGT]. Many other studies arrive at much lower estimates of the EIS—in some cases, substantially below one. *See, e.g.*, Tomas Havranek, Roman Horvath, Zuzana Irsova & Marek Rusnak, *Cross-Country Heterogeneity in Intertemporal Substitution*, 96 J. INT'L ECON. 100, 101-02 (2015) (finding a mean value of 0.5 in meta-analysis of 2,735 EIS estimates from 104 countries); Tomas Havranek, *Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting*, 113 J. EUR. ECON. ASS'N 1180, 1196 (2015) (estimating an EIS of 0.33 after correcting for publication bias in reported estimates). Importantly, Atkinson and Stiglitz's core claim regarding the suboptimality of capital taxation does not depend upon the EIS; however, if the EIS is less than one, then a capital tax will cause individuals to delay rather than accelerate consumption.

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

For Atkinson and Stiglitz, the key question thus becomes: which tax system raises the most revenue from high-θ individuals—a period-one labor income tax on its own, or a lower period-one labor income tax combined with a period-two capital tax?[41] The intuition for the former is that it still allows high-θ types to choose their preferred allocation between period-one and period-two consumption, without any distortion of their savings decision. The opportunity to choose one's own preferred allocation of period-one and period-two consumption is itself valuable, which suggests that high-wage earners should be willing to pay for that opportunity. Indeed, they should be willing to pay more to live in a capital-tax-free world *than they would have paid in period-two capital tax* (in present value terms), because the period-two capital tax not only costs them money but circumscribes their freedom of choice. The presence of a capital 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 tax tells us that there is an alternative policy involving only a labor income tax that would allow the government to extract even more from high-θ types, thus enabling even more redistribution and leaving 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.[42] 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 *controlling for period-one labor income*,[43] an additional tax on capital simply generates

41. Stiglitz elaborates on this point in Joseph E. Stiglitz, *Self-Selection and Pareto Efficient Taxation*, 17 J. PUB. ECON. 213, 236-38 (1982).

42. *See* Stiglitz, *supra* note 41, at 238.

43. 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*

an additional distortion without providing the government with any additional information about anyone's θ. A labor income tax is necessary even though it is distortionary because the tax system must rely on labor income for informational purposes. A capital tax, by contrast, results in distortion without a purpose.

3. Relationship to Consumption Taxation

Atkinson and Stiglitz's prescription for tax-free saving can be implemented in two ways. One is to exclude capital and capital income from tax. The other is to tax all income but to allow an immediate deduction for savings. The immediate-deduction yield-exemption equivalence—sometimes known as the Cary Brown theorem—connects the idea of tax-free saving with consumption taxation.

To illustrate: imagine first that an individual earns $100 in period one and that savings will grow at a 50% rate between period one and period two. The individual seeks to smooth consumption across the two periods (to consume the same amount in each). Now consider two possible tax regimes. In the first, labor income is taxed at a 50% rate, and capital income (the return from savings) is untaxed. The individual will therefore want to save $20 in period one, leaving $50 to pay the period-one tax and $30 for period-one consumption. The $20 in period-one savings will grow at a 50% rate to $30 in period two. 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 two. (Note that this setup—no deduction for savings but no taxation of capital income—aligns with Roth-style retirement accounts like Roth IRAs and Roth 401(k)s.[44])

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 one will be $100 minus $40 equals $60, and her period-one tax will be $30. Her $40 of savings and $30 of tax will leave $30 for period-one consumption. The $40 of period-one savings will grow at a 50% rate to $60 in period two. When she withdraws that $60 for consumption, she will pay a 50% tax ($30), leaving $30 for consumption in period two. (Note that this setup—an immediate deduction for savings with full taxation of dissavings—aligns with traditional-style retirement accounts.[45])

Non-Linear Income Taxation and Heterogeneous Tastes, 83 J. PUB. ECON. 217, 228 (2002) ("[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 also discussion *supra* Section I.A.1; and Bankman & Weisbach, *supra* note 39, at 1444-45.

44. *See* I.R.C. § 408A (Roth IRAs); § 402A (Roth 401(k)s).

45. *See* I.R.C. § 408 (traditional IRA); § 401(k).

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

	No Capital Income Taxation (Yield Exemption)	**Consumption Taxation (Immediate Deduction)**
Period One		
Labor income	$100	$100
Savings	$20	$40
Tax	50% x $100 = $50	50% x ($100 – $40) = $30
Consumption	$100 – $20 – $50 = $30	$100 – $40 – $30 = $30
Period Two		
Savings + growth (1.5x)	$20 x 1.5 = $30	$40 x 1.5 = $60
Tax	$0	50% x $60 = $30
Consumption	$30 – $0 = $30	$60 – $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: what one earns but doesn't save, one consumes. Thus, we can (at least for now) understand a consumption tax as equivalent to a zero rate of tax on capital income, as both leave the taxpayer in the same position. Likewise, we can understand Atkinson and Stiglitz's rationale for not taxing capital income as an argument for a consumption tax.

4. Relationship to Chamley-Judd

The Atkinson-Stiglitz zero-capital-tax result shares similarities with another staple of late twentieth-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.[46] Like the Atkinson-Stiglitz

46. Christophe Chamley, *Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives*, 54 ECONOMETRICA 607, 607-09 (1986); Kenneth L. Judd, *Redistributive Taxation in a Simple Perfect Foresight Model*, 28 J. PUB. ECON. 59, 59-61 (1985).

theorem, the Chamley-Judd result forms an element of the intellectual background to NDPF, though the influence of the Chamley-Judd result on legal scholarship has been more muted.[47]

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% capital income tax rate and a 10% interest rate, the additional commodity tax one period out is relatively small: instead of being able to substitute $1 of period-one consumption for $1.10 of period-two consumption, an individual can substitute $1 of period-one consumption for $1.08 of period-two consumption. Put differently, the capital income tax reduces the rate of return from 10% pre-tax to 8% after-tax. Or said yet another way, the capital income tax is equivalent to a commodity tax on period-two consumption of 1.8% ($0.02/$1.10 ≈ 1.8%).

Now consider how a 20% capital income tax with a 10% interest rate affects the tradeoff between consumption today and consumption in 200 years. Absent capital income taxation, $1 of present-period consumption grows at a 10% rate and can be substituted for $189,905,276 of consumption in 200 years ($1 x $1.10^{200} \approx$ $189,905,276). With a 20% capital income tax, $1 of present-period consumption grows at an 8% after-tax rate and can be substituted for only $4,838,950 of consumption two centuries from now ($1 x $1.08^{200} \approx$ $4,838,950). The 20% 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 (for example, 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 ultimately approach 100%. 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 and leaves less income available for redistribution.[48]

---

47. As of November 2025, fewer than twenty articles in the Westlaw database of legal secondary sources make any mention of the Chamley-Judd result. *See supra* note 20. It is not entirely clear why the Atkinson-Stiglitz theorem has been so much more influential than the Chamley-Judd result among legal scholars. One possible answer is that while Joseph Bankman and David Weisbach translated the Atkinson-Stiglitz theorem into terms that non-economists could easily understand in their 2006 *Stanford Law Review* article, *see generally* Bankman & Weisbach, *supra* note 39, no one has done the same for Chamley-Judd.

48. Unlike the Atkinson-Stiglitz theorem, the Chamley-Judd result does depend on whether the elasticity of intertemporal substitution is greater than or less than one. *See supra* note 40. For an explanation of why capital taxation may be desirable in the long run when the EIS is less than one, see Ludwig Straub & Iván Werning, *Positive Long-Run Capital Taxation: Chamley-Judd Revisited*, 110 AM. ECON. REV. 86, 88 (2020).

5. Policy Impact

Measuring the policy impact of first-generation optimal tax theory is difficult because a variety of forces shaped tax policy across advanced economies in the last quarter of the twentieth century and the first quarter of the twenty-first. "As a matter of fact," observe Thomas Piketty and Emmanuel Saez, "all advanced economies impose substantial capital taxes" notwithstanding the stark zero-capital-tax result emerging from the Atkinson-Stiglitz and Chamley-Judd models.[49] And indeed, Atkinson and Stiglitz themselves emphasized that their zero-capital-tax result was based on "strong assumptions,"[50] urging policymakers to consider factors—such as administrability—that their model excludes.[51]

Still, even though no country has followed the zero-capital-tax result to its logical endpoint, policymakers have certainly taken note of first-generation optimal tax theory's core insights. In the waning days of President Gerald Ford's administration, the Treasury Department cited a working-paper version of Atkinson and Stiglitz's seminar article in a report extolling the virtues of consumption taxation, which—as noted—is one method of excluding capital and capital income from the tax base.[52] Martin Feldstein, who would go on to serve as Chairman of the Council of Economic Advisers under President Reagan, also cited the Atkinson-Stiglitz zero-capital-tax result in advocating for capital tax reform.[53] More recently, Kevin Hassett, the Chairman of the Council of Economic Advisers in the first Trump administration and the Director of President Trump's National Economic Council as of this writing, has cited the Chamley-Judd model in testimony to Congress calling for capital tax cuts.[54]

In the meantime, capital income tax rates have plummeted in the United States and other advanced economies. According to the Congressional Research Service, the effective marginal tax rate on capital income in the United States—taking into account corporate-level and shareholder-level taxes—fell by half, from 46% to 23%, in the quarter century from

49. Thomas Piketty & Emmanuel Saez, *A Theory of Optimal Capital Taxation* 1 (Nat'l Bureau of Econ. Rsch., Working Paper 17989, Apr. 2012), https://eml.berkeley.edu/~saez/piketty-saezNBER12optKtax.pdf [https://perma.cc/F8GJ-YRWY].

50. See ANTHONY B. ATKINSON & JOSEPH E. STIGLITZ, LECTURES ON PUBLIC ECONOMICS 471 (2d ed. 2015). For example, the zero-capital-tax result depends upon the assumption that the government is free to set the labor income tax schedule optimally. *See id.* (noting that "[i]f . . . the government is not free to vary the nonlinear [labor] income tax schedule . . . , then we cannot appeal to the [zero-capital-tax] result").

51. *Id.* at 472-73.

52. *Blueprints for Basic Tax Reform*, U.S. DEP'T OF THE TREASURY 217 (Jan. 17, 1977), https://home.treasury.gov/system/files/131/Report-Blueprints-1977.pdf [https://perma.cc/QKV5-4G24].

53. *See* Martin Feldstein, *The Welfare Cost of Capital Income Taxation*, 86 J. POL. ECON. S29, S33 & n.12 (1978).

54. *How the Taxation of Capital Affects Growth and Employment: Hearing Before the Joint Econ. Comm.*, 112th Cong., 2d Sess. 293-40 (2012) (testimony of Dr. Kevin A. Hassett).

1978 to 2003.[55] Other advanced economies have undergone a similar transformation: according to one estimate, the average corporate-level plus shareholder-level tax rate across members of the Organisation for Economic Co-operation and Development fell from 75.2% in 1981 to 42.0% by 2012.[56] It would, of course, be a mistake to attribute all of that decline to the work of optimal tax theory. But optimal tax theory has provided an intellectual infrastructure for the transformation of tax policy across the developed world. As Piketty and Saez note, "the zero capital tax result remains an important reference point . . . in policy discussions."[57] Put another way, first-generation optimal tax theory has served as a lodestar—though far from the only driving force—for changes to capital taxation in the real world.

B. The New Dynamic Public Finance

As with the static Mirrleesian approach, a key figure in the emergence of NDPF is Mirrlees himself, whose 1978 paper co-authored with Peter Diamond—*A Model of Social Insurance with Variable Retirement*[58]—planted the seed for the new dynamic public finance (though the NDPF literature would not take off for another two decades). But unlike Mirrlees's more famous 1971 article, that 1978 paper and its many extensions are virtually unknown among legal scholars, cited only four times in the Westlaw database of legal secondary sources—twice by David Weisbach,[59] once by Daniel Shaviro,[60] and once by this author.[61]

The key move in the NDPF literature is to transform the one-shot model from Mirrlees's 1971 article into a repeated game in which individual θ levels and government policies both potentially undergo change. A typical NDPF model starts with a set of *self-reports* and *allocations*. All individuals self-report their θ level to the planner, and the planner allocates a certain amount of "leisure"[62] and consumption to each individual based

55. CONG. RESEARCH SERV., RS21706, HISTORICAL EFFECTIVE MARGINAL TAX RATES ON CAPITAL INCOME 12 tbl.1 (2006).

56. OECD Overall Dividend Tax Rates (Corporate and Personal), 1981-2012, TAX FOUND. (July 3, 2012), https://taxfoundation.org/blog/oecd-overall-dividend-tax-rates-corporate-and-personal-1981-2012 [https://perma.cc/2BGW-67J8].

57. *See* Piketty & Saez, *supra* note 49, at 1.

58. P.A. Diamond & J.A. Mirrlees, *A Model of Social Insurance with Variable Retirement*, 10 J. PUB. ECON. 295 (1978).

59. *See* David A. Weisbach, *Toward a New Approach to Disability Law*, 2009 U. CHI. LEGAL F. 47, 50 n.10 (discussing the 1978 paper in a single paragraph of a footnote); David A. Weisbach, *What Does Happiness Research Tell Us About Taxation?*, 37 J. LEGAL STUD. S293, S316 (2008) (citing the 1978 paper but noting that it is "not very helpful in the present context").

60. *See* Shaviro, *supra* note 16, at 785, n.113 (describing the 1978 Diamond-Mirrlees paper as "important" and devoting a paragraph to the dynamic Mirrleesian approach).

61. *See* Daniel J. Hemel, *Phaseouts*, 77 TAX L. REV. 53, 93-94 & n.153 (2023).

62. Economists typically use the term "leisure" to refer to all time not devoted to market labor. This definition results in, for example, the time devoted to caring for one's own children

on her self-reported θ level. The planner seeks the set of allocations that maximizes social welfare, subject to the constraint that individuals will self-report their θ level truthfully only if the bundle of leisure and consumption associated with that self-reported θ level is at least as attractive as the bundle associated with all lower self-reported θ levels. This "incentive-compatibility constraint" is the limiting factor on redistribution: the planner cannot make the low-θ bundles too attractive or else high-θ types will falsely report low-θ levels.[63]

In contrast to static optimal tax models, the "game" between self-reporting individuals and the planner is played over and over in NDPF models. Individuals' true θ values may change from one period to the next. Alternatively, individuals may self-report a high θ in one period and a low θ in the next period even if their true θ value doesn't change. And of particular significance, the *planner*'s strategy may change: a planner that offers one set of allocations in an earlier period may shift to offering a different set of allocations in a later period.

Along with self-reporting and allocation, another concept that is methodologically central to dynamic taxation is the idea of *wedges*. A wedge is any difference between the marginal rate of substitution and the marginal rate of transformation. The marginal rate of substitution refers to the rate at which an individual is willing to exchange one item for another. The marginal rate of transformation refers to the rate at which society can convert one item into another.[64]

For example, imagine a concert violinist whose labor—in the form of playing classical music—generates enjoyment for others. Let's say that listeners are willing to pay $1,000 for the joy of hearing the violinist play for an additional hour. Thus, when the concert violinist sacrifices an hour of leisure and devotes it to labor (here, music playing), society gains $1,000 worth of consumption (here, music listening). The marginal rate of transformation for the violinist is one hour of leisure (labor) for $1,000 of consumption. (Recall that the marginal rate of transformation refers to what the violinist produces, not what she receives in remuneration.)

Now imagine that if the concert violinist truthfully reports her θ level, the planner will allocate to her a bundle of leisure and consumption that leaves her indifferent between an additional hour of leisure and $600 of consumption. There would thus be a 40% *wedge* between the concert

and family members being defined as "leisure." *See* Daniel J. Hemel & David A. Weisbach, *The Behavioral Elasticity of Tax Revenue*, 13 J. LEGAL ANALYSIS 381, 386 (2021) (noting standard terminology and why it is confusing). I will use the term "leisure" in text to align with the terminology used by economists, recognizing that the term "leisure" is misdescriptive in important ways.

63. *See* Mikhail Golosov, Aleh Tsyvinski & Iván Werning, *The New Dynamic Public Finance: A User's Guide*, *in* 21 NBER MACROECON. ANN. 317, 325 (Daron Acemoglu, Kenneth Rogoff & Michael Woodford eds., 2007).

64. *See* Narayana R. Kocherlakota, *Wedges and Taxes*, 94 AM. ECON. REV. 109, 109 (2004).

violinist's marginal rate of substitution (1 hour for $600) and the marginal rate of transformation for the concert violinist's labor (1 hour for $1,000).

The wedge in the violinist example is known as a *labor wedge*. Another type of wedge is an *intertemporal consumption wedge*. To continue with an example from the discussion of the Chamley-Judd result in Section I.A.4, let's again say that the interest rate is 10%. When an individual sacrifices $1 of present consumption and instead lends out her $1, someone else can invest that $1 to generate goods or services worth $1.10 next year. Thus, the marginal rate of transformation is $1 of present consumption for $1.10 of next-year consumption. Now imagine that the planner allocates consumption across time such that a particular individual is indifferent between $1 of present consumption and $1.08 of consumption next year. That individual's marginal rate of substitution is $1 of present consumption for $1.08 of next-year consumption. The intertemporal consumption wedge is approximately 1.8% ($2/$110).

After specifying the welfare-maximizing set of allocations and corresponding wedges, the NDPF approach aims to find *implementations* of those allocations.[65] Economists working in the dynamic Mirrleesian tradition understand, of course, that individuals do not literally self-report their θ levels to a planner, and a planner does not literally allocate leisure and consumption across individuals. However, the NDPF literature highlights the fact that actual tax and non-tax policies generate all sorts of real-world wedges.[66] The goal at this last stage is to identify policies that can implement—or at least approximate—the wedges in the welfare-maximizing set of allocations.

In our concert violinist example, one way to implement a 40% labor wedge is to impose a 40% tax rate on the violinist's earnings. The violinist earns $1,000 per hour in ticket receipts pre-tax but ends up with only $600 per hour after tax, so she is willing to work only until she is indifferent between an hour of leisure and $600 of consumption. In our second example, one way to implement a 1.8% intertemporal consumption wedge is to impose a 20% tax on capital (interest) income. A 20% tax on capital income (with our assumed 10% interest rate) means that $1 of savings yields $1.08 of next-year consumption. The individual therefore saves until she is indifferent between $1 of present consumption and $1.08 of consumption next year.

The implementations of the wedges in the previous paragraph are relatively straightforward, but some implementations are less so. For example, a capital income tax typically creates both an intertemporal consumption wedge and a labor wedge. The intertemporal consumption wedge causes individuals to shift consumption from later periods to earlier periods. More consumption today, in turn, means that the marginal utility of

65. *See* Golosov, Tsyvinski & Werning, *supra* note 63, at 332.
66. *See id*. at 332-34.

current consumption declines. That decline in the marginal utility of consumption widens the labor wedge: an individual will be *less* willing to substitute leisure for current consumption because she derives less utility, at the margin, from current consumption. (This idea that capital income taxation generates a labor wedge lies at the heart of the Atkinson-Stiglitz model discussed in Section I.A.2.)

The effect of a capital income tax on labor wedges will depend on how much people save and over what time period. Assume that older workers who save now are planning to consume those savings relatively soon, while younger workers who save now are planning to consume those savings far in the future. As illustrated in the discussion of the Chamley-Judd result in Section I.A.4, a capital income tax assessed on an accrual basis each year imposes a heavier implicit tax on far-in-the-future consumption than on near-term consumption. Thus, a capital income tax at a constant rate (for example, 20%) will induce a larger shift in the intertemporal consumption choices of younger workers than of older workers. As younger workers shift more of their consumption to the present, the marginal utility of current consumption declines. Thus, the labor wedge for younger workers widens: they become less willing to trade leisure for current consumption. The upshot is that a constant-rate capital income tax—the same 20% rate for everyone—generates a wider labor wedge at younger ages than at older ages.[67] Or as Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger put it, "a positive capital income tax mimics a labor income tax that is falling with age."[68]

Importantly, taxes are not the only way to implement wedges. Long-term employment contracts with fixed wage rates likewise result in wedges: if an individual's productivity rises but her wage remains the same, then her labor wedge will widen (because the marginal rate of transformation from leisure to consumption rises while her marginal rate of substitution remains the same). Other sources of wedges include public insurance arrangements (for example, Social Security Disability Insurance[69] and unemployment insurance) and private insurance.[70]

Articles in the dynamic Mirrleesian tradition typically prescribe specific quantitative details of the implementations that they propose, and the math quickly gets quite complicated. However, the basic insights can be explained qualitatively. The remainder of this section turns toward those basic insights.

---

67. *See* Andrés Erosa & Martin Gervais, *Optimal Taxation in Life-Cycle Economies*, 105 J. ECON. THEORY 338, 340 (2002); Carlos Garriga, *Optimal Fiscal Policy in Overlapping Generations Models*, 47 PUB. FIN. REV. 3, 5 (2019).

68. Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger, *Taxing Capital? Not a Bad Idea After All!*, 99 AM. ECON. REV. 25, 41 (2009) (emphasis omitted).

69. *See* Mikhail Golosov & Aleh Tsyvinski, *supra* note 14, at 271.

70. *See* Jonathan Heathcote & Hitoshi Tsujiyama, *Optimal Income Taxation: Mirrlees Meets Ramsey*, 129 J. POL. ECON. 3141, 3142 (2021).

1. Changing Productivities

Much of the NDPF literature focuses on the effect of changing productivities (that is, changes in real output per hour worked). Recall that the 1971 Mirrlees model condensed time to a single period, and the Atkinson-Stiglitz model precluded the possibility of changing productivities because individuals were assumed to work only in the first period. One of NDPF's main contributions—partly attributable to Mirrlees himself[71]—is to show how optimal tax analysis is affected by the possibility of random productivity shocks over multiple periods of labor.

A simplified example will serve to illustrate. Imagine two periods, two possible θ levels (high and low), and two wage levels (high and low). Individuals who have high θ levels in period one may subsequently experience negative θ shocks, rendering them low θ in period two. (For expository ease, we will set aside the possibility that individuals with low θ levels in period one may experience positive θ shocks and become high θ in period two. For now, assume that high-θ types in period one can be high θ or low θ in period two but low-θ types in period one will remain low θ in period two.)

Imagine that the tax rate on high-wage earners in period one is set as high as it can go without inducing high-θ types to mimic. In period two, some of those period-one 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 so they can enjoy more leisure.

71. *See generally* Diamond & Mirrlees, *supra* note 58 (developing a foundational framework for optimal social insurance and retirement timing within the Mirrleesian tradition).

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

	Period-Two High Wage	**Period-Two Low Wage**
Period-One High Wage	Period One: High-θ / Period Two: High-θ "high-high"	Period One: High-θ / Period Two: High-θ "high-low" (mimicking)
		Period One: High-θ / Period Two: Low-θ "high-low" (negative θ shock)
Period-One Low Wage		Period One: Low-θ / Period Two: Low-θ "low-low"

Note that the "mimicking" high-lows in the top right box of Table 3 are not necessarily nefarious. They may be individuals who are high θ in period one and prudently accumulate savings to insure themselves against the possibility of a negative θ shock in period two. Because they do not know ex ante whether they will experience such a shock, they save more than they would have saved had they known they would remain high-θ types in period two. Once period two rolls around and these individuals do not experience a negative θ shock, they find themselves with an excess of savings. The excess of savings in period two reduces their marginal utility of consumption and makes them less willing to trade leisure for a high-wage job in period two. Thus, the term "mimicking" does not reflect a normative judgment. It simply signifies that a person is a high-θ type but behaves like a low-θ type.

The simplified setup in Table 3 sets the stage for analysis of three distinct but related issues: history dependence, capital income taxation, and age dependence.

History Dependence. One of the core questions addressed by 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 second-period 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 (a larger second-period transfer for high-lows than for low-lows). Increasing the transfer to

high-lows in period two will ease the incentive-compatibility constraint in period one: high-θ types in period one will have a stronger incentive to report truthfully (that is, to earn high wages) in period one because now, a high period-one wage guarantees a larger transfer in the event of a negative θ shock in period two. Because high-θ types in period one will value the greater protection from negative period-two θ shocks, they should be willing to pay extra for the additional insurance. Therefore, the period-one tax rate on high-wage earners can rise (relative to the previous revenue-maximizing rate) without inducing high-θ types to mimic low-θ types in period one. Even low-lows will be better off because now the government can extract more revenue from high-θ types in period one: assuming high-θ types are risk averse, high-θ types should be willing to pay more for the insurance than what it costs the government to provide.

But while a larger transfer to high-lows eases the incentive-compatibility constraint in period one, it tightens the incentive-compatibility constraint in period two. In the latter period, earning a low wage will be 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 one but harms incentives for high-θ types in period two. Past-income progressivity—that is, a smaller transfer to high-lows than to low-lows—does the opposite.

As a matter of theory, it is thus not obvious whether past-income regressivity or past-income progressivity will be optimal.[72] However, the case for past-income regressivity becomes stronger if the government can observe negative θ shocks. For example, if the government can observe which individuals experience debilitating injuries that render them incapable of work in period two, then the government may want to implement a period-two transfer conditional on disability that is regressive with respect to past income. Past-income regressivity in period two will ease the incentive-compatibility constraint for high-θ types in period one: revealing oneself to be a high-θ type in period one now results in greater insurance against a negative θ shock. Meanwhile, the larger transfer to verifiably disabled high-lows won't undermine incentives for genuine high-highs to reveal themselves truthfully in period two because high-highs can't access the disability-contingent transfer.

72. 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." *See* Marek Kapicka, *Quantifying the Welfare Gains from History Dependent Income Taxation* 3 (May 13, 2022) (unpublished manuscript), https://economie.esg.uqam.ca/wp-content/uploads/sites/54/2022/05/DOT_Main.pdf [https://perma.cc/FLA7-CAFC]. In other words, high-lows ought to pay less in period-two tax (or receive larger period-two transfers) than low-lows, which in turn allows for higher period-one taxes on high-wage earners. However, Kacpika acknowledges that the results are quite sensitive to the model inputs. *Id.* at 34–35.

Things change if the government cannot observe θ shocks. In that case, past-income regressivity (i.e., a larger transfer to high-lows in period two) undermines the incentive for high-highs to reveal themselves truthfully in period two because they now receive a larger transfer for not working or working in a low-wage job. Put another way, however tempting it was for high-θ types to mimic in period one, it will be even more tempting for them to mimic in period two if the transfer to self-reported high-lows rises. More generally, 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 (earning a high wage in period one and a low wage in period two) will be more attractive to high-θ types if they can save some of their period-one 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 taxes and other forms of capital 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 two.

Imposing a capital income tax in period two, 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 one. Or, put differently, it widens the period-one labor wedge. Capital income taxation thus encounters the same tradeoff as history-dependent taxation: changes that improve incentives for high-θ types in period two come at the expense of incentives for high-θ types in period one.

One potential resolution of this tradeoff is to target the capital income tax at high-lows. In other words, the capital income tax would be progressive over past income and regressive over current labor income, with individuals who earned more labor income in the last period and less in the current period facing higher capital income tax rates. Recall that the rationale for the capital income tax is to dissuade high-θ types from presenting themselves as low θ (in other words, to earn a low wage) in period two. The rationale for imposing the tax therefore does not apply to individuals who report low wages in period one or high wages in period two.[73]

What about high-lows who actually do experience negative θ shocks in period two? Perhaps counterintuitively, a capital income tax that is

73. 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, 1607-08 (2005) (finding that agents with unexpected high labor income in later periods optimally face negative marginal capital tax rates).

regressive over current labor income can leave these genuine high-lows better off. High-θ types who don't plan on mimicking in period two 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. The capital income tax thus serves to screen out mimickers from the population of high-lows (albeit imperfectly). That, in turn, means that the government can redistribute more to the high-lows without violating the incentive-compatibility constraint. Genuine high-lows benefit from screening out the mimickers.

Notably, the role for capital taxation envisioned by the NDPF literature—deterring high-θ types from mimicking low-θ types in period two—arises only because of the occurrence of negative productivity shocks. If productivity remained constant over the life cycle, then the government would know that every high-low is mimicking and would want to tax them the same as high-highs. But of course, in the real word, we know that some high-lows *aren't* mimicking—they experience genuine negative θ shocks (for example, because they become ill, because they have to care for a family member who has become ill, or because their own skills have obsolesced due to economic and technological change). The existence of genuine high-lows motivates the effort to screen out mimicking high-lows.

In sum, NDPF analysis points to a role for capital income taxation when individuals are vulnerable to negative productivity shocks. Under those circumstances, a capital income tax can ease the period-two incentive-compatibility constraint by deterring high-θ types from mimicking. Deterring high-θ mimickers allows the government to provide larger transfers to genuine high-lows. Thus, the capital income tax facilitates provision of a more complete θ-shock insurance product.

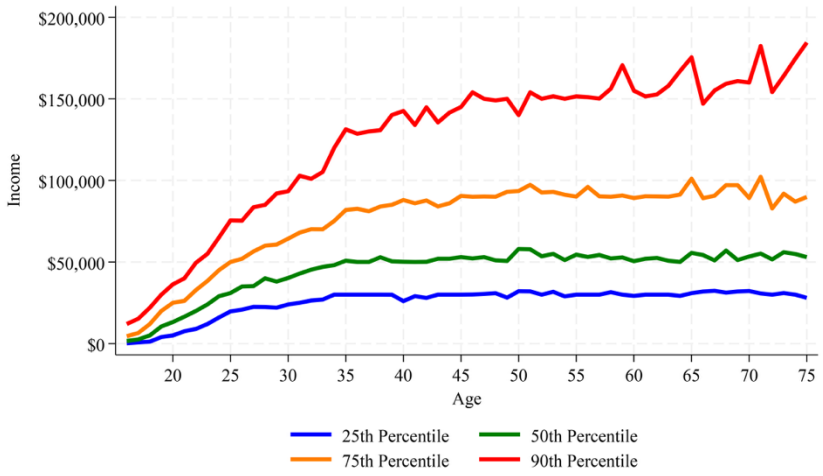
The Role of Age. Age plays a particularly important role in dynamic analyses of labor and capital income taxation. This is so for three reasons. First, as Figure 1 illustrates, income increases over the life cycle (and especially over the course of the twenties and thirties). If capital markets are perfect, individuals in their twenties and thirties will be able to borrow against their future income in order to smooth consumption. However, if capital markets are imperfect (for instance, 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.[74] Thus, one way to raise welfare is to shift consumption from the old to the young.

Second, and even more starkly, inequality increases over the life cycle. U.S. adults in their twenties have much lower incomes than their older

74. *See* Daniel Shaviro, *Permanent Income and the Annual Income Tax* 21-22 (N.Y.U. Law & Econ. Rsch. Paper Series, Working Paper No. 06-28, July 2006), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=920622 [https://perma.cc/97HA-7EL4].

counterparts, but the distribution of income among twenty-somethings is relatively egalitarian. 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 three times the public defender),[75] but the income gap between a Cravath partner and an experienced public defender in New York is far larger (the average Cravath partner makes roughly twenty times the highest-paid public defenders).[76]

Figure 1: Pre-Tax Income by Age and Percentile[77]



75. *Compare* Patrick Smith, *Cravath Tops Davis Polk Salary Scale for 4th Years and Above*, AM. LAW. (Feb. 28, 2022), https://www.law.com/americanlawyer/2022/02/28/cravath-tops-davis-polk-salary-scale-for-4th-years-and-above [https://perma.cc/ENA5-LWNF] (highlighting a $215,000 salary for Class of 2022 attorneys), *with* Jonah E. Bromwich, *Hundreds Have Left N.Y. Public Defender Offices Over Low Pay*, N.Y. TIMES (June 9, 2022), https://www.nytimes.com/2022/06/09/nyregion/nyc-public-defenders-pay.html [https://perma.cc/GMP5-HFGH] (noting a $74,000 starting salary for first-year criminal public defenders at Legal Aid Society of New York).

76. *Compare Cravath*, LAW.COM, https://www.law.com/compass/firm/LF00000273/Cravath/overview [https://perma.cc/PA6A-5YYV] (noting Cravath profits per partner of $6,050,000 in 2023), *with Legal Aid Society*, PROPUBLICA: NONPROFIT EXPLORER, https://projects.propublica.org/nonprofits/organizations/135562265 [https://perma.cc/F6Y9-E8NH] (noting total compensation of $325,703 for top-paid Legal Aid Society attorney in fiscal year 2023).

77. Figure 1 is based on U.S. Census Bureau, *Current Population Survey, 2020 Annual Social and Economic (ASEC) Supplement* (2020). Each person in the ASEC Supplement sample who was 15 years or older was asked to report "the amount of money income received in the preceding calendar year" from eleven sources, including money wages or salary, net income from self-employment, Social Security, public assistance or welfare payments, interest, dividends, unemployment and workers' compensation, pensions, alimony, and child support. Income was reported on a pre-tax basis. *See Current Population Survey, 2020 Annual Social and Economic (ASEC) Supplement*, U.S. CENSUS BUREAU 7-3 (2020), https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar20.pdf [https://perma.cc/7KQP-HNLL].

Third, moral hazard increases as workers approach retirement age. Relatively few high-θ individuals in their fifties leave the labor market in order to enjoy more leisure. Lots of high-θ individuals in their early to mid-sixties do. One recent study estimates that among males, the extensive margin labor supply elasticity in the age sixty-one to sixty-five group is twelve times as high as in the age fifty-one to sixty group.[78] Thus, the cost of providing θ-shock insurance is much higher for individuals in their sixties than for individuals in their thirties, forties, and fifties. (The extensive margin labor supply elasticity is also very high for individuals in their twenties—and particularly for women in that age demographic.[79])

The combination of these three factors—increasing income, increasing inequality, and increasing moral hazard—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 life cycle 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 twenty-somethings still would pay positive taxes and very low-income fifty-somethings still would receive net transfers.

Second, from behind the veil of ignorance, we would want to buy more θ-shock insurance for our fifties than for our twenties and thirties because we face much greater uncertainty about income in our fifties than in our twenties and thirties. Progressive income taxes are a form of θ-shock insurance, so this observation suggests that tax progressivity should rise from early adulthood into one's fifties.

Third, we would likely want to buy less θ-shock insurance for our sixties than for our fifties. 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 fifty as a result of rising intra-cohort inequality and then fall in later life as a result of rising moral hazard.[80]

If tax rates cannot be conditioned explicitly on age, then age-independent capital subsidies can partially implement the desired age-

78. See Marios Karabarbounis, *A Road Map for Efficiently Taxing Heterogeneous Agents*, 8 AM. J. ECON.: MACROECON. 182, 200 tbl.3 (2016) (estimating Frisch elasticity of labor supply for males 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*. (estimating Frisch elasticity of labor supply for females to be 0.81 in age 61-65 group and 0.18 in age 51-60).

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

80. *See* Karabarbounis, *supra* note 78, at 183, 203 fig.4, 207 fig.6; Abdoulaye Ndiaye, *Flexible Retirement and Optimal Taxation* 44-45 (N.Y.U. Stern Sch. of Bus., Rsch. Paper No. 1, Aug. 14, 2020), https://papers.ssrn.com/sol3/Delivery.cfm/5262721.pdf?abstractid=5262721 [https://perma.cc/NU7S-CM29].

dependent labor income tax schedule. Recall from our discussion of the Chamley-Judd result that an age-independent capital income tax creates wider labor wedges for younger individuals than for older individuals because younger individuals typically save for a more distant future.[81] Symmetrically, age-independent capital *subsidies* narrow the labor wedge for younger individuals relative to older individuals because younger individuals have more years in which to accumulate those subsidies. Thus, one might favor age-independent capital subsidies for the same reason that one might want income taxes to rise with age: because the need for θ-shock insurance rises with age. Note, though, that age-independent capital subsidies can only implement labor wedges that rise with age. They cannot replicate the right-side drop of the hump-shaped schedule suggested by dynamic tax analysis.

Importantly, there is 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 (such as the child tax credit,[82] the earned income tax credit,[83] and the additional standard deduction for the aged[84]) already depend on age. One insight from the NDPF literature is that when we impose age-invariant taxes, we are throwing out a huge amount of information that could be used to provide superior θ-shock insurance products.

2. Changing Policies

Along with considering changes in productivity, the NDPF literature also wrestles with the problem of policy change. Anyone who lived through the early 1990s in the United States knows that "read my lips, no new taxes" does not necessarily mean that there will be no new taxes: governments can change their tax-and-transfer policies regardless of what they promise.[85] A common challenge for governments is to persuade their populations that promises about future policy are credible. The NDPF literature formalizes and advances the study of this challenge.

The problem is clearest in the capital tax context. Consider again the Atkinson-Stiglitz model: the government taxes labor income in period one but promises not to tax capital in period two.[86] The NDPF literature poses the question: what's to prevent the government from reneging on its

81. *See supra* notes 67-68 and accompanying text.

82. I.R.C. § 24(c)(1), (i)(2)-(3).

83. I.R.C. § 32(c)(1)(ii)(II).

84. I.R.C. § 63(c)(3).

85. *See* Lily Rothman, *The Story Behind George H.W. Bush's Famous 'Read My Lips, No New Taxes' Promise*, TIME (Dec. 1, 2018), https://time.com/3649511/george-hw-bush-quote-read-my-lips [https://perma.cc/Q4DC-2LZW].

86. *See supra* notes 39-40 and accompanying text.

commitment, confiscating all savings in period two, and redistributing those savings equally? A confiscatory capital tax in period two would improve welfare, because it would shift resources from richer individuals (who have a low marginal utility of consumption) to poorer individuals (who have a high marginal utility of consumption). And if it were truly one-time, then it would not distort labor and savings decisions.[87] One-time confiscation is "free" redistribution ("free" in the sense that it generates no deadweight loss).[88]

The potential problem with a large one-time tax on capital is, of course, that individuals might expect it to happen again.[89] Fear of confiscation then might induce high-θ types to supply less labor and to decumulate capital. Not only is this outcome inefficient,[90] but it also leaves low-θ types worse off because the government will struggle to raise revenue, resulting in less redistribution.

The solution might seem to be for the government to commit not to confiscate capital. But what makes that commitment credible? As Stanley Fischer emphasized in an important 1980 article, even a *benevolent* government will have an incentive to violate its prior commitments.[91] In the moment, confiscation is welfare-maximizing! On top of that, the fact that the government refrains from confiscating wealth today does not mean it will refrain tomorrow. If individuals expect confiscation to occur sometime in the future—and if they modify their labor and savings behavior accordingly—then nonconfiscation today is a missed opportunity to improve welfare.

Unlike the Atkinson-Stiglitz model, which simply assumes that governments can make credible commitments, the NDPF literature—consistent with its game-theoretic origins—takes the problem of credible commitment seriously.[92] It makes three significant contributions to the study of time inconsistency.

First, as the late Emmanuel Farhi and coauthors emphasize, nontaxation of capital today can actually *increase* the risk of confiscation—and,

87. Chamley and Judd both acknowledge this point. *See* Chamley, *supra* note 46, at 619; Judd, *supra* note 46, at 60.

88. A large one-time capital tax might actually encourage labor effort insofar as it would induce high-θ types who had planned to play the high-low strategy to opt for high-high instead because their savings would have been confiscated.

89. On the problem of time inconsistency in fiscal policy, see generally Finn E. Kydland & Edward C. Prescott, *Rules Rather than Discretion: The Inconsistency of Optimal Plans*, 85 J. POL. ECON. 473 (1977).

90. *See* Kevin Roberts, *The Theoretical Limits to Redistribution*, 51 REV. ECON. STUD. 177, 192 (1984).

91. *See* Stanley Fischer, *Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government*, 2 J. ECON. DYNAMICS & CONTROL 93, 100-01 (1980).

92. On the importance of credible commitment to game theory, see generally JOHN VON NEUMANN, ON THE THEORY OF GAMES OF STRATEGY (1928), *reprinted in* 4 CONTRIBUTIONS TO THE THEORY OF GAMES 13 (Albert William Tucker & Robert Duncan Luce eds., 1959).

reciprocally, taxation of capital can reduce confiscation risk.[93] Farhi et al. proceed from the assumption that a confiscatory capital tax will entail some current costs—for example, the fixed costs of enacting and implementing the tax, along with potential domestic and international reputational repercussions. The costs are worth bearing only if the benefits are large enough. If the distribution of capital across society were relatively egalitarian, the redistributive gains from confiscation likely wouldn't be worth the costs. By adopting a progressive capital tax, the government can ensure that wealth inequality never exceeds the threshold at which confiscation becomes 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 tax of inequality"[94]—the threat of expropriation that arises from a lopsided distribution of resources. One might view this as the formal version of FDR's intuition that a robust government response to the Great Depression could prevent the rise of socialism in the United States.[95] 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.[96] 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 way the balance ultimately

93. Emmanuel Farhi, Christopher Sleet, Iván Werning & Sevin Yeltekin, *Non-linear Capital Taxation Without Commitment*, 79 REV. ECON. STUD. 1469, 1469-70 (2012).

94. *See* Daniel J. Hemel, *Capital Taxation in the Middle of History*, 99 N.Y.U. L. REV. 1554, 1562 (2024).

95. *See* President Franklin D. Roosevelt, Address Accepting the Presidential Nomination at the Democratic National Convention in Chicago (July 2, 1932), https://www.presidency.ucsb.edu/documents/address-accepting-the-presidential-nomination-the-democratic-national-convention-chicago-1 [https://perma.cc/5KTZ-DKHF] ("To meet by reaction that danger of radicalism is to invite disaster. . . . The way to meet that danger is to offer a workable program of reconstruction . . . ."); President Franklin D. Roosevelt, Address at the Democratic State Convention, Syracuse, N.Y. (Sep. 29, 1936), https://www.presidency.ucsb.edu/documents/address-the-democratic-state-convention-syracuse-ny [https://perma.cc/3JKK-M35E] ("We were against revolution. Therefore, we waged war against those conditions which make revolutions—against the inequalities and resentments which breed them."); *see also* SEYMOUR MARTIN LIPSET & GARY MARKS, IT DIDN'T HAPPEN HERE: WHY SOCIALISM FAILED IN THE UNITED STATES 73-74 (2000) (arguing that Roosevelt made "conscious efforts to undercut left-wing radicals, to preserve capitalism," by—among other measures—"advancing tax reforms designed to stop an unjust concentration of wealth and economic power" (internal quotation marks omitted)).

96. Alberto Bisin & Adriano A. Rampini, *Markets as Beneficial Constraints on the Government*, 90 J. PUB. ECON. 601, 603-04 (2006).

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.

Although Bisin and Rampini's argument might seem inconsistent with Farhi et al.'s, a common thread connects the two: governments can bolster their non-confiscation commitments either by reducing their *incentive* to confiscate (Farhi et al.) or their *capacity* to confiscate (Bisin and Rampini). As we will see in Parts II and III, this insight will have important real-world implications. For example, progressive capital taxation (which reduces the government's incentive to expropriate) and deregulation of cryptocurrency (which reduces the government's ability to expropriate) can be understood as substitute strategies for addressing the time-inconsistency problem.

Third, while the literature on time inconsistency in fiscal policy typically focuses on capital accumulation, NDPF also highlights the relationship between government credibility and *labor* incentives. Lack of credible commitment distorts labor incentives not only because it casts a shadow capital tax (which in turn generates a labor wedge), 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 2021 to 2025. Nothing we do in 2026 or thereafter will affect our tax liability—in that sense, the tax will be like a lump-sum tax. This retrospective labor income tax would be a highly efficient way of redistributing from high-θ types to low-θ types: past income isn't a perfect proxy for θ, 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. High-θ types would have an additional reason to mimic low-θ types in the present: to avoid high retrospective labor income taxes in the future.

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 applies to anonymous labor markets. Limiting the amount of information available to the government about our earnings histories means that individuals can earn high wages today without fear that the government will impose a retrospective tax on today's wages down the road.

* * *

Summing up so far, the new dynamic public finance challenges and enriches the "old" public finance in several ways. By extending the basic Mirrlees model to a world with multiple periods of labor and random θ shocks, NDPF highlights the welfare-improving potential of history-dependent taxes, capital income taxes, and age-dependent taxes. By taking seriously the problem of credible commitment, NDPF draws attention to the ways in which time inconsistency threatens the government's ability to redistribute—as well as ways in which the government can make its own commitments more credible. While these insights may seem abstract at present, they will have concrete implications for real-world policy choices. The next parts turn toward those implications.

II. Implications for Taxes and Transfers

In some cases, the NDPF approach helps to rationalize previously puzzling features of our existing tax-and-transfer system. In other cases, NDPF offers arguments for reform rather than justifications of the status quo. This Part focuses on NDPF's implications for five features of the U.S. tax-and-transfer system: the pattern of history dependence embedded in social insurance programs (Section II.A); the existence of positive capital income taxes notwithstanding the Atkinson-Stiglitz zero-capital-tax prescription (Section II.B); the age invariance of the federal income tax schedule (Section II.C); the separate rate schedules for single individuals and married couples (Section II.D); and the absence of a value-added tax in the United States (Section II.E).

A. 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. One example of a history-dependent transfer program is Social Security Disability Insurance (SSDI). Monthly SSDI benefits are based on average lifetime earnings pre-disability,[97] so 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% of an individual's pre-unemployment wages up to a maximum amount (for example, $869 per week in New York as of this writing).[98] Thus, "high-lows" receive larger transfers than "low-lows" under UI.

97. *See* SOC. SEC. ADMIN., PUB. NO. 05–10029, DISABILITY BENEFITS 9 (Feb. 2025), https://www.ssa.gov/pubs/EN-05-10029.pdf [https://perma.cc/4JVP-Z8C9].

98. *What is the Maximum Benefit Rate?*, N.Y. STATE DEP'T OF LABOR, https://dol.ny.gov/mbr [https://perma.cc/3BHW-E3FV].

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.[99]

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."[100]

NDPF 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 one is that it entitles you to a larger transfer if you become disabled or unemployed in period two. The historical dependence of SSDI and UI benefits thus relaxes the incentive-compatibility constraint in period one, potentially allowing the government to impose higher labor income taxes without inducing mimicking.

To be sure, these improved period-one incentives come at the cost of period-two incentives, since a larger period-two 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-two 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-two wage and relaxing on the beach in one's spare time. If high-θ types are risk averse and the government can verify the negative θ shock (disability or involuntary termination), then the value of the additional insurance to high-θ types who earn high wages in period one 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 one, 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-two moral hazard for high-θ types. Without verification, by contrast, it is theoretically ambiguous whether the positive effect of past-income regressivity on period-one incentives for high-θ types outweighs the negative effect on period-two incentives (inducing true high-highs to mimic high-lows).

At times, the Internal Revenue Code has incorporated history dependence too. The Revenue Act of 1964 implemented a form of "income

99. Robert E. Goodin, *Compensation and Redistribution*, 33 NOMOS 143, 169 n.18 (1991).
100. *Id*. at 148.

averaging": individuals whose income in a given year exceeded 133.33% of their average income over the previous four years received relief according to a formula designed to ensure that they wouldn't pay more in taxes than if their income had been spread evenly over the five-year period.[101] Significantly, income averaging under the 1964 Act applied only to low-highs (individuals whose incomes went *up*), not to high-lows (individuals whose incomes went *down*). While it would have required extraordinary prescience for Congress to anticipate the insights of NDPF years in advance, Congress's decision to exclude high-lows from income averaging under the 1964 Act accords with NDPF logic: without some mechanism to verify that individuals who report high incomes in one period and low incomes in the next have in fact experienced a negative θ shock, income averaging for high-lows would potentially induce mimicking by "true" high-highs.[102] In any event, the 1964 Act's general income-averaging provisions were repealed as part of the 1986 tax reform, and today, only a select group of taxpayers (individuals engaged in farming and fishing businesses) have the option of income averaging.[103]

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 verifiable or semi-verifiable negative θ shocks. It also points to a reason why the federal income-tax system—which lacks a similar mechanism to verify negative θ shocks—might therefore eschew history dependence in the case of high-lows. This is not to imply that current SSDI and UI benefit schemes are optimal in all respects (indeed, they almost certainly are not).[104] 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 overall progressivity and historical independence of general income taxation—may not be so incongruous after all.

101. Revenue Act of 1964, Pub. L. No. 882-72, § 232, 78 Stat. 19, 105.

102. For a proposal to extend income averaging to high-lows, see generally Lily L. Batchelder, *Taxing the Poor: Income Averaging Reconsidered*, 40 HARV. J. ON LEGIS. 395 (2003). The 2007 article by Daniel Shaviro mentioned in the introduction succinctly summarizes the NDPF case against high-low income averaging. *See* Shaviro, *supra* note 16, at 777.

103. *See* Taxpayer Relief Act of 1997, Pub. L. No. 105-34, § 933(a), 111 Stat. 788, 881 (codified as amended at I.R.C. § 1301) (allowing taxpayers engaged in farming businesses to average their income over a four-year period if their income increases); *see also* American Jobs Creation Act of 2004, Pub. L. No. 1083-57, § 314(b), 118 Stat. 1418, 1468 (amending I.R.C. § 1301) (extending income averaging to taxpayers engaged in fishing businesses). These new income averaging provisions for farming and fishing businesses, like their predecessor, apply to taxpayers whose income increases—in other words, for low-highs but not for high-lows.

104. On flaws in the design of UI systems, see generally Brian Galle, *How to Save Unemployment Insurance*, 50 ARIZ. ST. L.J. 1009 (2018).

B. Capital Taxation

The NDPF literature's prescription for capital income tax rates that are regressive over current labor income may seem at odds with the federal statutory rate structure, which effectively imposes capital income tax rates that *rise* with current labor income. (Capital income tax rates—tax rates on interest, dividends, and capital gains—depend on taxable income, which in turn encompasses both capital and labor income. Thus, the tax rate on capital income is generally higher for someone with high labor income than for someone with low labor income.[105]) 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 investment income over a certain inflation-adjusted threshold ($11,950 in tax year 2025).[106] The investment income limit functions as an astronomically high marginal tax rate on the 11,951st dollar of capital income for individuals with low labor income (an 804,600% rate for individuals claiming the maximum credit for tax year 2025).[107] A second example is Supplemental Security Income (SSI), which is available only to individuals with "countable resources" of $2,000 or less ($3,000 for a couple). Countable resources include stocks and bonds, bank accounts, motor vehicles, and real estate.[108] A similar asset test applies to some Medicaid beneficiaries.[109] 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 (though again, there is no evidence that policymakers were aware of NDPF's insights when they designed these programs). The implicit capital taxes in these transfer programs 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

105. See I.R.C. § 1(j) (detailing a rate schedule in effect for tax years 2018 through 2025).

106. I.R.C. § 32(j); *see* Rev. Proc. 2024-40, § 2.06(2), 2024-45 I.R.B 1100.

107. *See* Rev. Proc. 2024-40, § 2.06(2), 2024-45 I.R.B 1100 (establishing a maximum credit of $8,046 in 2025 for taxpayers with three or more qualifying children).

108. With exemptions for one vehicle and one personal residence. *Spotlight on Resources -- 2025 Edition*, SOC. SEC. ADMIN., https://www.ssa.gov/ssi/spotlights/spot-resources.htm [https://perma.cc/72YY-R39Z].

109. 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 sixty-five and over. *See New York Medicaid Eligibility for Long Term Care: Income & Asset Limits*, AM. COUNCIL ON AGING (Feb. 24, 2025). https://www.medicaidplanningassistance.org/medicaid-eligibility-new-york [https://perma.cc/Y9X2-QS5H].

existing EITC investment income limit and the SSI and Medicaid asset limits still may be flawed.[110] However, insofar as these features are flawed, it is not because they are regressive over current labor income. Capital taxes that are regressive over current labor income can play a useful role in achieving a more redistributive tax-and-transfer system overall.

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 tax-favored retirement accounts (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 between the time of contribution and the time of withdrawal, cash-flow taxation produces a positive effective tax rate on savings, and when marginal rates fall, cash-flow taxation generates a savings subsidy. 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 other words, a capital subsidy).

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—exactly the people on whom NDPF suggests we should impose positive capital taxes. The traditional IRA/401(k) structure facilitates consumption smoothing by would-be mimickers, which is exactly what an NDPF-inspired approach would seek to prevent.

The NDPF literature thus sheds light on the long-running contest between traditional and Roth retirement accounts.[111] In particular, NDPF insights provide a strong reason for policymakers to favor Roth plans over traditional plans. Traditional IRAs and 401(k) plans subsidize saving by the very individuals whose savings we should discourage: individuals who earn high incomes in "period one" (before they turn 59 ½ and become

110. For example, recent research by Felix Wellhmied suggests that optimal asset-test thresholds are significantly higher than those in SSI and Medicaid. *See* Felix Wellschmied, *The Welfare Effects of Asset Mean-Testing Income Support*, 12 QUANTITATIVE ECON. 217, 240 (2021).

111. *See, e.g.*, William G. Gale, J. Mark Iwry & Gordon McDonald, *An Analysis of the Roth 401(k)*, 110 TAX NOTES 163, 165-67 (2006) (criticizing Congress's decision to allow taxpayers to opt for Roth treatment of 401(k) defined contribution plans).

eligible for tax-free IRA withdrawals) and earn low incomes in "period two" (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 life cycle. Roth IRAs and 401(k) plans, by contrast, subsidize the savings of low-highs: individuals who earn low incomes earlier in life and high incomes later on. While there may be no persuasive reason to subsidize savings for low-highs either, saving behavior by low-highs is less likely to implicate mimicking concerns than saving by high-lows.

C. Age Dependence

Statutory income tax rates do not depend explicitly 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 life cycle means that average rates will typically rise for taxpayers into middle age, in accordance with NDPF prescriptions.[112] Likewise—though the magnitude of the effect is much smaller—the additional $1,600 standard deduction for individuals sixty-five and older[113] slightly reduces average rates for senior citizens, in accordance with the 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 twenty-five to sixty-four.[114] 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 rates on labor income should be lower for young adults and seniors than for middle-aged 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.[115] 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.

---

112. *See* Martin Gervais, *On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System*, 36 J. ECON. DYNAMICS & CONTROL 682, 682 (2012).

113. I.R.C. § 63(f); *see* Rev. Proc. 2024-40, § 2.15(3), 2024-45 I.R.B 1100.

114. I.R.C. § 32(c)(1)(A)(ii)(II). These age restrictions were temporarily suspended for tax year 2021. *See id*. § 32(n)(2).

115. For an important—and prescient—exception, see generally Fennell & Stark, *supra* note 18.

D. Marriage

One of the most salient features of the U.S. federal income tax system is the existence of separate rate schedules for single and married taxpayers. Economists have only begun to extend dynamic Mirrleesian insights to the tax treatment of marriage,[116] but this emerging line of research suggests a possible rationale for applying different rate schedules to single and married taxpayers. As Mikhail Golosov and Ilia Krasikov observe, married individuals are less vulnerable to θ shocks than single individuals because married individuals have access to a private θ-shock insurance mechanism: intra-marital redistribution.[117] Since the optimal amount of θ-shock insurance depends upon the degree to which individuals are exposed to θ shocks, Golosov and Krasikov's observation suggests that the government should provide less θ-shock insurance—or in other words, set a less progressive tax schedule—for married individuals than for single individuals.

Reinforcing Golosov and Krasikov's observation is a second reason why the government might differentiate between single and married individuals in terms of the amount of θ-shock insurance it provides: different degrees of moral hazard. For both men and women, marriage is correlated with a higher price elasticity of labor supply[118]—or put another way, married people appear to be more sensitive to labor income tax rates. Correlation, of course, does not necessarily imply causation: it is possible that people with higher labor supply elasticities are more likely to get married, rather than marriage leading to higher labor supply elasticities.[119] Either way, insofar as married people are more responsive to tax rates, the government has an additional reason to provide less θ-shock insurance to married individuals than to single individuals: not only are married individuals less exposed to the risk of θ shocks, but they are also costlier to insure because the moral hazard of θ-shock insurance is greater.

To be sure, nothing in the NDPF literature suggests that the current tax treatment of single and married individuals in the United States is anywhere close to optimal. Under current law, married couples who file joint returns are taxed on their combined income: the Internal Revenue Code does not differentiate between a dollar earned by one spouse versus the other. Golosov and Krasikov note that the government may be able to

116. Mikhail Golosov & Ilia Krasikov, *The Optimal Taxation of Couples*, 140 Q. J. ECON 2163, 2190 (2025) (analyzing optimal nonlinear taxation for couples and showing the relevance of multi-dimensional mechanism design to household tax policy).

117. *See id.* at 2165.

118. *See* Todd Elder, Steven J. Haider & Cody Orr, *The Evolution of the Wage Elasticity of Labor Supply Over Time* 38 tbl.2 (IZA Inst. of Labor Econ., Working Paper No. 16393, Aug. 2023), https://docs.iza.org/dp16393.pdf [https://perma.cc/2DQP-MKFC].

119. For evidence that the relationship between marriage and the price elasticity of labor supply among women may reflect causation and not merely correlation, see Olivier Bargain, Libertad González, Claire Keane & Berkay Özcan, *Female Labor Supply and Divorce: New Evidence from Ireland*, 56 EUR. ECON. REV. 1675, 1682 (2012).

increase social welfare by applying different tax schedules to "primary" and "secondary" earners—in other words, higher- and lower-earning members, respectively, of two-worker couples.[120] One reason to differentiate between primary and secondary earners is that the labor supply of secondary earners tends to be more elastic.[121] Put another way, moral hazard is greater for secondary earners than for primary earners—which potentially justifies lower tax rates for secondary earners because taxes are more likely to distort secondary earners' labor-supply choices.[122]

Summing up: the same logic that leads to greater progressivity for middle-aged taxpayers than for younger adults in NDPF models—increased exposure to negative θ shocks—also supports greater progressivity for single individuals than for married people. Bolstering the case for differential taxation of single individuals and married couples—and for greater progressivity with respect to the former—is the fact that married couples are more susceptible to moral hazard. As in the case of history dependence and capital taxation, NDPF helps to rationalize one aspect of the U.S. tax-and-transfer system: the application of different rate schedules to single individuals and married couples. At the same time, NDPF suggests possible reforms that could improve upon the status quo, such as separate rate schedules for primary and secondary earners within married couples.

E. 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 NDPF supplies arguments in *favor* of capital income taxation. One might therefore think that NDPF provides an argument *against* VATs—a suggestion that the United States, almost 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.

First, 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

120. *See* Golosov & Krasikov, *supra* note 116, at 40 & 11 n.13.

121. See Robert McClelland, Shannon Mok & Kevin Pierce, *Labor Force Participation Elasticities of Women and Secondary Earners Within Married Couples* 30 tbl.7 (Cong. Budget Off., Working Paper 2014-06, Sep. 2014), https://www.cbo.gov/publication/49433 [https://perma.cc/ZP6D-VA6J].

122. For further discussion, see Daniel Hemel, *Beyond the Marriage Tax Trilemma*, 54 WAKE FOREST L. REV. 661, 681-690 (2019). For a brief period in the early to mid-1980s, Congress did allow married couples to deduct a percentage of the lower-earning spouse's income, thus effectively reducing the rate on secondary earners. On the legislative history of the short-lived secondary earner deduction, see *id.* at 690 & n.103.

a VAT entails a one-time tax on existing wealth plus zero net capital taxation going forward.[123] For example, when the government imposes a 20% 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.[124]

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 defy easy explanation,[125] 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,[126] and their VATs have not triggered discernible capital flight.) Concededly, it is possible that Americans will see a VAT as a strategy to impose a one-time capital levy (which, indeed, a VAT is). 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—the government simply receives revenue from vendors at various stages of the supply chain. 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.[127] This feature of a VAT will be particularly attractive to governments that otherwise would not be able to make credible non-exploitation commitments. Note, though, that the anonymity argument for a VAT envisions a total transition from personal income taxation to impersonal value-added taxation. Layering a VAT on top of a personal income tax—which most other countries have done—would not achieve the same anonymity benefit.

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 the NDPF literature

123. *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).

124. *Id*.

125. For further discussion, see Hemel, *supra* note 94, at 1619-23.

126. *See* ORG. FOR ECON. COOP. & DEV., CONSUMPTION TAX TRENDS 2020: VAT/GST AND EXCISE RATES, TRENDS AND POLICY ISSUES 265-69 (2022), https://www.oecd.org/en/publications/consumption-tax-trends-2022_6525a942-en.html [https://perma.cc/EZ8B-A4UG].

127. 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.

can justify many puzzling features of the U.S. tax-and-transfer system, the absence of a VAT in the United States is not one of them.

* * *

The implications of NDPF for the tax-and-transfer system may seem scattershot at first glance. But they all connect to the two themes running throughout NDPF analysis: changing productivities and changing policies. The motivation for history dependence is that a history-dependent tax-and-transfer schedule can better distinguish true productivity changes from feigned ones. The motivation for tax-and-transfer schedules that depend upon age and marriage status is that those variables affect individuals' risk of and vulnerability to productivity shocks—and a well-designed tax-and-transfer system is, to a large extent, a society-scale productivity shock insurance program. Explicit capital taxation can supplement the screening function of history dependence and—through its effect on inequality—can mitigate the threat of sudden changes in policy (for example, a one-time capital levy). By adding the possibility of productivity changes and policy changes to first-generation optimal tax models, NDPF opens up new vistas for both descriptive and normative analysis of the tax-and-transfer landscape.

Table 4: New Dynamic Public Finance Prescriptions and U.S. Tax-and-Transfer Policy

	New Dynamic Public Finance	**U.S. Tax-and-Transfer Policy**
History Dependence	Labor income taxes should be regressive on past income when negative θ shocks are verifiable	Potentially supports: Disability insurance and unemployment insurance benefit formulas
Capital Taxation	Capital taxes should be positive and regressive on past income	Potentially supports: Positive tax rates on capital income Potentially supports: Investment income limit for EITC Does not support: Traditional IRAs/401ks
Age Dependence	Labor income taxes should rise with age until one's 60s and then fall	Potentially supports: Larger standard deduction for taxpayers age 65 and older Does not support: Disallowance of EITC for childless individuals under age 25 and for those age 65 or older
Marriage	Labor income taxes should be more progressive for single individuals than for married couples	Potentially supports: Different tax schedules for single and married filers
Value Added Tax	VAT allows one-time capital tax without stoking fear of reprise	Does not support: Absence of a value added tax in the US

III. Implications for Other Areas of Law

While NDPF primarily focuses on tax-and-transfer policies, its implications extend well beyond the tax-and-transfer domain. This final part draws lessons from NDPF for non-tax areas of law—including torts, contracts, property, securities regulation, rules around cryptocurrency, and constitutional law.

A. Torts

A central question in tort law is whether—and to what extent—the tort system should seek to insure individuals against accidental losses.[128] Although the insurance function played an important role in mid-twentieth-century tort-law theory,[129] subsequent scholars have adopted a more skeptical stance toward the insurance rationale for torts. As the late Yale law-and-economics theorist George Priest put it in a much-cited 1987 lecture: "Tort law is an extremely perverse method of providing compensation insurance to consumers . . . ."[130] According to Priest, a first-party insurance system would be far more efficient than tort law as a means of shielding individuals against negative income shocks, and thus "our society would benefit if the insurance features of modern tort law were excised."[131] Instead of insuring individuals against accidental losses, tort law should—according to Priest—focus primarily or exclusively on deterrence (i.e., on reducing the number of accidents in the first place).[132]

An NDFP perspective potentially casts the insurance function of tort law in a more favorable light. From an NDPF vantage point, asymmetric information is a fundamental problem for the tax-and-transfer system: government officials struggle to verify whether—and to what extent—individuals have experienced negative θ shocks. That information asymmetry between the government and putative victims of those negative shocks is what allows true high-highs to masquerade as high-lows. One way to bridge this information gap might be to implement a system through which claims of negative θ shocks could be tested—perhaps through an adversarial

128. *See* Beatrice A. Beltran, Comment, *Posner and Tort Law as Insurance*, 7 CONN. INS. L.J. 153, 156 (2000) (noting that the relationship between tort law and the insurance function "has spurred enormous debate").

129. *See* Fleming James Jr., *Social Insurance and Tort Liability: The Problem of Alternative Remedies*, 27 N.Y.U. L. REV. 537, 540 (1952) (observing that "something of the philosophy of social insurance has crept into the thinking about tort liability"); John C. P. Goldberg, *Twentieth-Century Tort Theory*, 91 GEO. L.J. 513, 537-44 (2003) (discussing the role of insurance in the mid-century "enterprise liability theory" of torts).

130. George L. Priest, *The Monsanto Lectures: Modern Tort Law and Its Reform*, 22 VAL. U. L. REV. 1, 20 (1987).

131. *Id*.

132. *See id*. at 20-22. For a critique of the tort-law-as-insurance idea from a corrective justice perspective, see generally Ernest J. Weinrib, *The Insurance Justification and Private Law*, 14 J. LEGAL STUD. 681 (1985). For another classic argument against the tort-law-as-insurance idea, see generally Jane Stapleton, *Tort, Insurance and Ideology*, 58 MOD. L. REV. 820 (1995).

process in which a devil's advocate is incentivized to contest the putative θ-shock victims' claims. Such a system would enable the state to accomplish more redistribution without inducing more mimicking.

Tort law is precisely such a system. If accident victims exaggerate their damages in tort actions, or if those damages result from the victims' own contributory or comparative negligence, tort defendants have strong incentives to push back against the victims' compensatory claims. The discovery process, trial, and jury proceedings all serve as means of verifying a victim's negative θ-shock report. These verification steps, in turn, reduce the likelihood that true high-highs will be able to masquerade as high-lows by falsely asserting that they have suffered injuries or by overstating their losses.

One way to think about tort law, then, is as an adjunct to the tax-and-transfer system with a series of robust verification mechanisms added on. The higher prices we pay for goods and services due to the tort system operate like a tax, and damages awards for lost wages function like a transfer targeted at individuals who can show that they have experienced negative θ shocks.[133] Tort law still suffers from disadvantages relative to an ideal insurance scheme—tort law may "over-insure" in some circumstances (e.g., by providing large pain-and-suffering awards in cases where individuals would not have chosen to purchase pain-and-suffering insurance[134]) and "underinsure" in other cases (e.g., by providing nothing at all to victims whose injurers are judgment-proof and to victims who experience negative θ shocks resulting from non-tortious causes). But what tort law does—and other social insurance schemes generally do not—is to incentivize private parties to come forward with information when claimants who haven't suffered from negative θ shocks assert that they have (or when claimants exaggerate the extent of their losses).

The NDPF perspective thus offers a twist on the familiar idea that plaintiffs' attorneys in tort actions function as "private attorneys general" who advance the social goal of deterrence as well as their own clients' interest in compensation.[135] NDPF points us toward the fact that defense lawyers in tort cases also serve a social function: to ensure that the only plaintiffs who can successfully claim damages for lost income are those who have experienced genuine negative θ shocks. The policing function of defense lawyers in tort actions operates to screen out genuine high-highs who are mimicking high-lows. Defense lawyers thus play a vital role with

133. *See* George L. Priest, *The Current Insurance Crisis and Modern Tort Law*, 96 YALE L.J. 1521, 1560 (1987).

134. For canonical critiques of pain-and-suffering damages on insurance grounds, see Alan Schwartz, *Proposals for Product Liability Reform: A Theoretical Synthesis*, 97 YALE L.J. 353, 362-67 (1988); and Robert Cooter, *Towards a Market in Unmatured Tort Claims*, 75 VA. L. REV. 383, 392 (1989).

135. *See, e.g.*, William B. Rubenstein, *On What a "Private Attorney General" Is—and Why It Matters*, 57 VAND. L. REV. 2129, 2152-53 & tbl. 2, 2172 (2004).

respect to tort law's insurance function, much like plaintiffs' attorneys advance tort law's deterrence function.

Viewing tort law in this light also helps us make normative sense of otherwise puzzling aspects of the tort system. Consider the curious fact that successful tort plaintiffs can recover for lost wages—with higher earners therefore receiving larger damages awards than lower earners. Several scholars—including Ronen Avraham[136] and the late Stephen Sugarman[137]—have criticized this aspect of current practice as "regressive." But if we understand tort law as a form of social insurance operating in parallel to Social Security Disability Insurance (SSDI) and unemployment insurance (UI), then it is perhaps less surprising—and maybe even less objectionable—that tort law would be "past-income regressive" in much the same way that SSDI and UI are. If plaintiffs who obtain large lost-income awards are genuine high-lows rather than mimicking high-lows, tort law can provide those awards without introducing a significant second-period moral hazard. And at least in theory, the larger transfer to genuine high-lows incentivizes high-θ types in period one to report truthfully (i.e., earn high wages) because one potential benefit of a high period-one wage is greater insurance against a period-two tort injury.[138]

Thus, when the tort system is analyzed in tandem with the tax system, the notion that tort damages are regressive becomes more nuanced. The opportunity to collect larger damages in the event of an injury is an additional benefit of earning a higher wage in period one. The additional benefit potentially eases the incentive-compatibility constraint with respect to high-θ types in period one, thereby allowing the government to further increase the tax rate on high-wage earners. And precisely because tort law—like SSDI and UI, but unlike the income tax system—incorporates a mechanism for verifying negative θ shocks, the concern that tort law will induce genuine high-highs to mimic high-lows in period two is at least somewhat attenuated.

Concededly, it may be possible for a non-tort insurance system—say, explicit public insurance or first-party private insurance—to replicate tort law's verification element. The government or a private insurer could, for example, offer bounties to whistleblowers who come forward with credible evidence showing that a putative victim has overclaimed. But tort law has an important advantage over newfangled whistleblower mechanisms:

136. *See* Ronen Avraham, *Putting a Price on Pain-and-Suffering Damages: A Critique of the Current Approaches and a Preliminary Proposal for Change*, 100 NW. U. L. REV. 87, 114-15 (2006).

137. *See* Stephen D. Sugarman, *Tort Reform Through Damages Law Reform: An American Perspective*, 27 SYDNEY L. REV. 507, 517 (2005).

138. Whether the past-income regressivity of tort damages actually has any effect on pre-injury work incentives is, to be sure, an empirical question. For a discussion of the circumstances in which the distribution of benefits and costs under non-tax law may generate deadweight loss, see Daniel J. Hemel, *Wealth, Schmealth, Welfare, and Schmelfare*, 59 WAKE FOREST L. REV. 1103, 1149-50 (2024).

centuries of precedent and practice. What tort law offers is a time-tested way for the state to verify negative θ shocks and to screen out mimickers from genuine high-lows.

B. Contracts

An important question in contract law is whether courts should be allowed to order specific performance of personal service agreements. While Anglo-American law bars specific performance in the labor-contract context,[139] scholars of law and economics have recently begun to question the rationales for that longstanding rule.[140] Missing from this debate, however, is any consideration of how specific performance of personal service agreements interacts with the tax-and-transfer system. The new dynamic public finance spotlights these interactions and potentially generates a forward-looking justification for the hoary Anglo-American rule.

The rule against specific performance of personal service agreements makes it more difficult for workers to shift income from periods in which their marginal tax rate is higher to periods in which their marginal tax rate is lower. Suppose, for example, that a thirty-year-old—seeking to "game" an age-dependent tax schedule—agrees with her employer that she will work for ten years but receive most of her compensation in years one, two, and three (when her age-dependent tax rate remains relatively low) instead of years eight, nine, and ten (when her age-dependent tax rate is much higher). If the employee quits in year four, the employer will have limited recourse. In some states (though not, significantly, in California[141]), the employer may be able to enforce a noncompete agreement that temporarily prevents the employee from working elsewhere.[142] 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-

139. *See* Lumley v. Wagner (1852) 42 Eng. Rep. 687 (Ch.) (noting that court cannot compel mezzo-soprano singer to perform at Her Majesty's Theatre).

140. *See* Kimberly D. Krawiec & Nathan B. Oman, *The Case for Specific Performance of Personal Service Contracts*, 110 IOWA L. REV. 751, 757 (2025) (arguing that "the per se rule against the specific enforcement of personal service agreements should be dispensed with"); D. A. Jeremy Telman, *When Specific Performance of Personal Service Contracts Is the Right Remedy*, 110 IOWA L. REV. ONLINE 117, 134 (2025) (agreeing with Krawiec and Oman that specific performance of personal service contracts should not be banned in all cases, but arguing that the remedy should be available only in "limited circumstances when the court can be very confident that either the order will facilitate settlement []or that both parties will perform the contract in good faith").

141. On the origins of California's prohibition on noncompetes, see Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete*, 74 N.Y.U. L. REV. 575, 613-19 (1999).

142. A Federal Trade Commission rule banning non-competes nationwide was struck down by the district court on August 20, 2024, and although the FTC appealed to the Fifth Circuit, the appeal has been held in abeyance and was reported to have been dismissed as of September 8, 2025. *See* Ryan LLC v. FTC, No. 3:24-CV-00986-E, 2024 WL 3879954 (N.D. Tex. Aug. 20, 2024); *see also* Ryan LLC v. FTC, No. 24-10951, (5th Cir. Sep. 8, 2025) (order dismissing appeal).

dependent and age-dependent taxes, as employers will be reluctant to pay employees up front for future performance if those contracts are unenforceable. In this instance, the common law of contracts is complementary to NDPF-inspired taxation.

While the rule against specific performance of employment contracts makes it hard for workers to shift income earlier in time, the common law of contracts enables income shifting in reverse. Imagine, for example, that a fifty-five-year-old and her employer agree that she will work for the employer for ten more years and receive most of her compensation in year ten (when her age-dependent tax rate has declined). Under these circumstances, the employer's commitment to pay a large year-ten 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: when tax rates decline over the life cycle, the common law of contracts facilitates gaming. Note, though, that the Internal Revenue Code already has provisions designed to prevent workers and their employers from deferring compensation for tax purposes.[143] The combination of contract-law limitations on income acceleration and tax-law restrictions on deferred compensation potentially facilitate NDPF-inspired rate schedules that vary over the life cycle.

To be sure, NDPF cannot justify a per se rule against specific performance of personal service agreements unless workers face tax rates that increase over time. However, even in the absence of an explicitly age-dependent tax schedule, marginal and average tax rates tend to increase over the first part of a worker's career because the tax system is progressive and earnings generally rise into middle age.[144] As a result, concerns about cross-period income-shifting apply even when tax schedules are facially age-independent and history-independent. Yet the new dynamic public finance, by highlighting the potential social benefits of age-dependent and history-dependent tax schedules, offers a new justification for the rule against specific performance of labor contracts at a time when scholars are increasingly questioning that rule's historic bases.

C. Property, Securities, and Cryptocurrency

Beyond tort law and contract law, NDPF analysis yields potentially important implications for a third common-law subject: property law. One of property law's central concerns is the "legibility" of holdings, where legibility refers to the extent to which ownership and value are susceptible to

143. *See* I.R.C. § 83(a); Minor v. United States, 772 F.2d 1472, 1474-76 (9th Cir. 1985).
144. *See supra* Figure 1.

observation and taxation.[145] As the late political scientist and anthropologist James C. Scott argued, the desire for legibility played a significant role in the rise of the freehold tenure system across early modern Europe.[146] 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.[147]

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: the legibility of property holdings increases what we characterized above as the "shadow tax" of expropriation since legibility makes the sudden seizure of property possible.

The tradeoff between the benefits and costs of legibility manifests in many contexts. For example, a centuries-old question in property law is whether to require that all conveyances of land be recorded in a central registry. England's 1677 Statute of Frauds—requiring conveyances to be written but not centrally recorded—represented a compromise between the Crown, which supported a central registry, and landowners who feared that a central registry would facilitate confiscatory taxation.[148] One might think of the Statute of Frauds question as a clash between the Farhi et al. view,[149] which posits that capital taxation can reduce inequality and therefore reduce the fear of confiscation, and the Bisin and Rampini view,[150] which suggests that anonymous markets also alleviate confiscatory fears. These are two radically different strategies for reducing the shadow tax of potential expropriation: one involves taxing capital; the other involves disabling the government from taxing capital.

A similar issue arises with respect to nonregistered bearer bonds. Most bonds are registered: ownership is centrally recorded, and transfer of ownership requires an update to the central list. With bearer bonds, by contrast, possession is presumptive evidence of ownership. Historically,

145. *See* David A. Weisbach & Daniel J. Hemel, *The Legal Envelope Theorem*, 102 B.U. L. REV. 449, 451 (2022).

146. JAMES C. SCOTT, SEEING LIKE A STATE: HOW CERTAIN SCHEMES TO IMPROVE THE HUMAN CONDITION HAVE FAILED 39-40 (Yale Univ. Press, 1998).

147. On De Soto's land-titling efforts, see generally HERNANDO DE SOTO, THE MYSTERY OF CAPITAL: WHY CAPITALISM TRIUMPHS IN THE WEST AND FAILS EVERYWHERE ELSE (2000).

148. On the run-up to the enactment of the Statute of Frauds, see generally Philip Hamburger, *The Conveyancing Purposes of the Statute of Frauds*, 27 AM. J. LEG. HIST. 354 (1983). *See also* Weisbach & Hemel, *supra* note 145, at 503-06 (framing the Statute of Frauds debate in legibility terms).

149. *See* Farhi et al., *supra* note 93.

150. *See* Bisin & Rampini, *supra* note 96.

bearer bonds have been used as a tool for tax evasion because they allow owners to hide their holdings from authorities.[151]

For a government pursuing the Farhi et al. strategy (i.e., reducing fear of confiscation by keeping inequality in check), bearer bonds are an obstacle insofar as they allow owners to hide their holdings from authorities and escape capital taxation. Consistent with this view, Congress in 1982 required U.S. Treasury bonds to be registered, and it barred issuers of bearer bonds from claiming income tax deductions for interest payments.[152] (The 1982 Act also denied tax-exempt status to bearer bonds issued by state and local governments—a provision that prompted two Supreme Court cases and ultimately resulted in the provision being upheld.[153]) However, if the Farhi et al. strategy fails (i.e., if capital taxation can't keep inequality sufficiently in check to counter fears of confiscation), then the opposite approach—allowing bearer bonds—may be optimal.

This same tradeoff now arises with respect to the regulation of cryptocurrency. Cryptocurrencies such as Bitcoin potentially allow individuals to hide their holdings from the government and thereby escape capital taxation.[154] This is certainly problematic for countries pursuing the Farhi et al. strategy, because the ease of evasion limits the capacity of capital taxation to reduce inequality. On the other hand, for countries pursuing the Bisin and Rampini strategy, cryptocurrency may bolster the credibility of the non-confiscation commitment: if the government can't observe and access the wealth of its citizens, then it will be more difficult for the government to seize that wealth.

The analysis here does not resolve the question of whether governments ought to pursue the Farhi et al. strategy or the Bisin and Rampini strategy. The answer potentially depends on the strength of other institutions and the degree of existing inequality, among other factors.[155] The key point for now is that NDPF's insights regarding capital taxation and credible commitment can clarify tradeoffs in areas far afield from tax, even if it does not tell us how those tradeoffs should ultimately be resolved.

D. Constitutional Law

Finally, the NDPF literature yields important implications for constitutional law. The clearest example is the debate over wealth taxation.

151. *See* South Carolina v. Baker, 485 U.S. 505, 508-09 (1988).

152. *See* Tax Equity and Fiscal Responsibility Act of 1982, Pub. L. No. 972-48, § 310, 96 Stat. 324, 595.

153. *See* South Carolina v. Regan, 465 U.S. 367, 370 (1984) (allowing the state to proceed with its challenge to the constitutionality of the provision denying tax-exempt status to state-issued bearer bonds); *Baker*, 485 U.S. at 505 (upholding the provision).

154. *See* Omri Marian, *Are Cryptocurrencies 'Super' Tax Havens?*, 112 MICH. L. REV. FIRST IMPRESSIONS 38, 39 (2013).

155. On the determinants of credible commitment in capital taxation, see generally Hemel, *supra* note 94.

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.[156] Scholars who argue that wealth taxes shouldn't be classified as "direct taxes" emphasize the importance of giving Congress "flexibility in meeting twenty-first century challenges."[157]

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 *in*flexibility. A constitutional rule that freely allows Congress to impose capital income taxes but severely restricts wealth taxes would make sense if we think that (a) capital taxation is desirable in moderation but (b) fear of limitless capital taxation may lead to large distortions. A wealth tax potentially allows the government to confiscate up to 100% of an individual's holdings in one fell swoop. By contrast, a capital income tax—even at a 100% rate—allows the government to seize only the nominal return on investment each year (historically 5% to 11% in advanced economies since 1870,[158] and potentially less in the low-interest-rate environment of the foreseeable future[159]). While over time, even a modest capital income tax can significantly reduce the value of holdings (as emphasized in the discussion of the Chamley-Judd result in Section I.A.4), a rule that limits the government to *income* taxation and precludes *wealth* taxation can potentially reduce fears of sudden expropriation. For a government that seeks to combat inequality through capital taxation while also keeping fears of expropriation in check, a constitutional rule that allows income taxation but precludes wealth taxation may be sensible.

The NDPF literature also bears implications for a related debate over retrospective taxation. In a trio of late 1920s decisions—*Nichols v. Coolidge*,[160] *Blodgett v. Holden*,[161] and *Untermyer v. Anderson*[162]—the Supreme Court held that certain tax provisions could not be applied to pre-

156. 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 sixteen times the rate in Connecticut. *See* Daniel Hemel, *How to Tax Wealth Constitutionally*, MEDIUM (Jan. 28, 2019), https://medium.com/whatever-source-derived/how-to-tax-wealth-constitutionally-863ce992ac7e [https://perma.cc/6LJZ-QGQH]; Alex Zhang, *The Wealth Tax: Apportionment, Federalism, and Constitutionality*, 23 U. PA. J.L. & SOC. CHANGE 269, 283 tbl.1 (2020); John R. Brooks & David Gamage, *The Wealth Tax Debate: Lessons from U.S. History*, 72 NAT'L TAX J. 269, 273-76 (2019).

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

158. *See* Òscar Jordà, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick & Alan M. Taylor, *The Rate of Return on Everything, 1870-2015*, 134 Q. J. ECON. 1225, 1241 tbl.II (2019). On the wealth-tax rate cap implicit in capital income taxes, see generally Daniel Hemel, *Taxing Wealth in an Uncertain World*, 72 NAT'L TAX J. 755 (2019).

159. *See* Lukasz Rachel & Lawrence H. Summers, *On Secular Stagnation in the Industrialized World*, 50 BROOKINGS PAPERS ON ECON. ACTIVITY 1, 45-48 (2019).

160. 274 U.S. 531 (1927).

161. 275 U.S. 142 (1927).

162. 276 U.S. 440 (1928).

enactment transactions. The precedential weight of the cases is now quite uncertain, but the Court has never overruled them explicitly.[163] *Nichols*, *Blodgett*, and *Untermyer* 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: if individuals expect that the government will impose retrospective labor income taxes in the future, they may be more reluctant to reveal themselves as high-θ (those with an ability to earn a high income) now. A constitutional prohibition on retrospective labor income taxation potentially mitigates that concern. 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.[164] However, NDPF helps us understand why flexibility in tax policy might not always be a virtue.

NDPF's implications for constitutional law do not stop at the constitutional law of tax. Consider, for example, the Supreme Court's three-factor framework for evaluating procedural due process claims under the Fifth and Fourteenth Amendments. The three-factor approach originated in *Mathews v. Eldridge*, in which the Court considered whether procedural due process prohibits the Social Security Administration from terminating an individual's disability benefits before holding a pre-termination hearing.[165] The Court in *Mathews v. Eldridge* famously stated:

> [I]dentification of the specific dictates of due process generally requires consideration of three distinct factors: First, the private interest that will be affected by the official action; second, the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or substitute procedural safeguards; and finally, the Government's interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.[166]

Translating *Mathews v. Eldridge* into NDPF terminology, we can understand the second factor as asking how likely it is that the Social Security

163. *See* United States v. Carlton, 512 U.S. 26, 34 (1994) (stating that *Nichols*, *Blodgett*, and *Untermyer* "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").

164. *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.").

165. *See* Mathews v. Eldridge, 424 U.S. 319, 323 (1976).

166. *See id*. at 335. On the "ever-expanding application of the *Mathews* balancing test" since the late 1970s, see Andrew Blair-Stanek, Twombly *Is the Logical Extension of the* Mathews v. Eldridge *Test to Discovery*, 62 FLA. L. REV. 1, 4, 8-12 (2010).

Administration will incorrectly classify a true high-low as a mimicker if it forgoes a pre-termination hearing. And we can interpret the third factor as asking—among other things—how else the government might be able to go about distinguishing true high-lows from mimickers.

The new dynamic public finance sheds light on both of these questions. We saw above that true high-highs are less likely to mimic high-lows when capital taxation limits their ability to carry savings from earlier periods to later ones.[167] Thus, when capital taxes are high, the fraction of true high-highs among the population of individuals claiming disability benefits will be smaller, and the risk of erroneous deprivation will be larger. Moreover, while the third *Mathews v. Eldridge* factor focuses on other *procedural* requirements to protect the fisc from mimickers, NDPF shows us that other *substantive* policies (e.g., capital taxes) can serve a similar function. NDPF thus teaches us that when capital taxes are in place, the case for robust procedural due process protections is stronger (and the government's interest in cutting off disability benefits without a pre-termination hearing is weaker) because the concern about true high-highs mimicking high-lows is less acute. And likewise, when robust procedural due process protections are in place, the case for capital taxation is stronger insofar as those protections limit the government's ability to weed out mimickers through administrative verification of negative θ shocks. Put another way, in the context of disability insurance and other θ shock-dependent social insurance programs, capital taxation and procedural due process protections are complements: more of one bolsters the argument for more of the other. This insight bears not only theoretical but also potential doctrinal relevance insofar as *Mathews v. Eldridge* requires courts to evaluate procedural due process through a cost-benefit prism. The capital tax regime—by influencing the ratio of true high-lows to mimickers—very much affects the costs and benefits of procedural due process protection.

* * *

As with its takeaways for the tax-and-transfer system, NDPF's implications for other areas of law are variegated. But they consistently relate back to the two themes—changing productivities and changing policies—that define NDPF's field of focus. Personal injury damages in tort are one of the legal system's primary responses to productivity shocks. Long-term labor contracts are one of the principal market mechanisms for managing productivity fluctuations. Meanwhile, property law, securities law, and cryptocurrency regulation all can operate as constraints on sudden policy changes. Constitutional law connects to both themes: to changing productivities insofar as procedural due process protections affect the government's ability to implement productivity-shock insurance programs; to

167. *See supra* Section II.B.

changing policies insofar as constitutional provisions constrain the legislature's ability to carry out sudden capital tax swings.

Seen in this light, the notion that developments in optimal tax theory might bear implications for areas of law ranging from land registries to noncompete clauses becomes less surprising. The risk of a productivity shock—illness, disability, or a change in economic conditions that results in unemployment or underemployment—is one of the main preoccupations in most of our lives and, naturally, a chief concern for many areas of law. The risk of a sudden policy change looms overhead every time we make a long-term investment of human, financial, or physical capital. The new dynamic public finance—as the study of productivity changes, policy changes, and policy responses to the threat of both—offers a promising framework for uniting the analysis of similar challenges across diverse domains.

Conclusion

Like other joinders 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. For example, law and the new dynamic public finance helps to explain why seemingly regressive features of Social Security Disability Insurance and unemployment insurance can facilitate *more* redistribution. Law and the new dynamic public finance also helps to explain why a society might want to impose high capital taxes on low-wage earners—a feature that we observe in the Earned Income Tax Credit, the Supplemental Security Income program, and Medicaid, but for which we have lacked a robust justification thus far. Weighing in the other direction, law and the new dynamic public finance casts serious doubt on the wisdom of existing federal tax subsidies for retirement savings—and, in particular, the structure of traditional IRAs and 401(k)s. And law and the new dynamic public finance highlights new dimensions to age-old debates regarding the insurance function of tort law, specific performance of labor contracts, and legibility in property law—while also shedding light on cutting-edge questions such as the regulation of cryptocurrency.

The conclusions emerging from law and the new dynamic public finance do not fit neatly within existing ideological categories. Although the overall goal of creating a more comprehensive system of redistribution and social insurance aligns with progressive priorities, law and the new dynamic public finance potentially justifies features of current law that, viewed in isolation, appear to be regressive (such as the structure of SSDI and UI benefit formulas). Ultimately, a serious effort to address the problem of economic inequality over the life cycle will require policymakers to consider solutions that, on first glance, appear counterintuitive. NDPF logic

certainly generates its fair share of counterintuitive conclusions—though, as we have seen, some of these counterintuitive conclusions already are reflected in existing law.

Looking to the future, three promising pathways lie ahead for law and the new dynamic public finance. The first, and most straightforward, applies NDPF insights to a broader range of problems in tax-and-transfer policy. As an illustration, NDPF bears potentially important implications for tax law's realization requirement, which generally provides that gains are not taxed until an asset is sold or exchanged.[168] As compared to a mark-to-market system that taxes economic gains as they accrue, the realization requirement delivers a subsidy to taxpayers that increases with the length of their holding period.[169] The subsidy has important implications for the NDPF model of optimal taxation over the life cycle. For example, the realization subsidy reduces the labor and intertemporal consumption wedges for younger savers relative to middle-age savers because younger savers can expect to have longer holding periods (insofar as they have more years of life before retirement and death). The realization requirement thus adds another element of age dependency to the tax system—one that roughly tracks the NDPF prescription for lower tax rates on younger adults than for adults in middle age.[170] This is not to say that the realization requirement is the optimal means of implementing an age-dependent tax schedule—or anywhere close. It is to say, though, that the life cycle effects of the current realization requirement are among the many relevant factors to consider in evaluating more comprehensive capital tax reforms.[171]

A second promising pathway applies NDPF insights to a wider array of topics beyond the tax-and-transfer domain. As Peter Diamond and James Mirrlees observed in their 1978 paper that inspired much of the later work in NDPF, "a general feature of moral hazard situations" is that "the provider of insurance would usually want to control trade in related commodities."[172] Diamond and Mirrlees offer the example of fire insurance companies wanting their clients to acquire fire extinguishers.[173] The government, as a θ-shock insurer, similarly might want to encourage individuals to reduce θ risk through other channels, which may—in turn—have important implications for non-tax legal arrangements. For example, individual workers are less exposed to θ shocks when they are organized into firms than when they operate as independent contractors. Indeed, one

168. *See* Dhammika Dharmapala & Daniel Hemel, The Realization Doctrine and the Optimal Taxation of Capital Income (Oct. 2024) (unpublished manuscript) (on file with author).

169. *See* David M. Schizer, *Realization as Subsidy*, 73 N.Y.U. L. REV. 1549, 1563 (1998).

170. *See supra* Figure 1 and text accompanying nn. 75-76.

171. For related work analyzing the realization requirement from an NDPF-informed perspective, see generally Mark Aguiar, Benjamin Moll & Florian Scheuer, *Putting the "Finance" into "Public Finance": A Theory of Capital Gains Taxation* (Nat'l Bureau of Econ. Rsch., Working Paper 32951, Sep. 2024), https://www.nber.org/papers/w32951 [https://perma.cc/3WHZ-BG73].

172. *See* Diamond & Mirrlees, *supra* note 58, at 304.

173. *See id.*

important function of the firm is to pool multiple risks (including health risk, longevity risk, mortality risk, disability risk, and productivity risk) across workers.[174] An NDPF-informed perspective may support employment-law changes that favor employer-employee relationships over independent contractor arrangements on the theory that firms are to θ-shock insurance what fire extinguishers are to fire insurance.

A final avenue for research connects law and the new dynamic public finance to one of its most recent "law and" predecessors: law and behavioral economics.[175] The new dynamic public finance typically conceives of individuals as well-informed actors who respond rationally to the incentives generated by tax-and-transfer rules. Behavioral economics has, of course, challenged that rational actor assumption,[176] and real-world policy prescriptions must respond to that challenge. For example, if individuals are overly sanguine about their risk of a negative θ shock—what the behavioral economics literature refers to as "optimism bias"[177]—then larger transfers to high-lows in period two are less likely to motivate high-θ types in period one (because overly optimistic high-θ types in period one do not expect to become low-θ types down the road). Importantly, behavioral biases do not negate the case for incorporating NDPF insights into law and public policy. Rather, the merger of behavioral economics with the new dynamic public finance shows how the wisdom of certain legal rules and practices—such as the past-income regressivity of disability insurance payments and tort damages—depends on how closely individuals hew to the rational actor model. Insights from NDPF can enable policymakers to translate behavioral "anomalies"[178] into prescriptions for legal reform.

In sum, law and the new dynamic public finance provides an underutilized toolkit for legal scholars and policymakers grappling with our era's "defining challenge."[179] While the highly technical nature of the NDPF economics literature imposes significant entry barriers, this article has sought to lower those barriers and make key insights from the literature accessible to a broader audience. These insights should, at the very least, lead scholars toward a clearer understanding of the ways in which tax and non-tax policies can mitigate—or exacerbate—economic inequality and

174. *See* Daniel J. Hemel, *Pooling and Unpooling in the Uber Economy*, 2017 U. CHI. LEGAL F. 265, 268-280 (2017).

175. For a recent survey of the law and behavioral economics literature, see EYAL ZAMIR & DORON TEICHMAN, BEHAVIORAL LAW AND ECONOMICS 1 (2018).

176. On behavioral economics' challenge to the rational-actor assumption, see generally Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CALIF. L. REV. 1051 (2000).

177. *See* Christine Jolls, *Bounded Rationality, Behavioral Economics, and the Law, in* 1 THE OXFORD HANDBOOK OF LAW AND ECONOMICS: METHODOLOGY AND CONCEPTS 60, 62-64, 67 (Francesco Parisi ed., 2017).

178. For an early description of some of these anomalies, see generally Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, *Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias*, 5 J. ECON. PERSP. 193 (1991).

179. Press Release, White House Office of the Press Sec'y, *supra* note 1.

insecurity over time. More optimistically, though less immediately, NDPF may inform the design of laws and policies that produce a more efficient and effective system of social insurance across the life cycle.