TAXATION AND INNOVATION—A SECTORIAL APPROACH

*David Hasen**

A number of tax rules have been adopted or proposed to promote innovation. The primary justification for these rules is that they can be effective in reducing or eliminating chronic market failure in the innovation sector. This Article argues that special tax rules for innovation generally are inappropriate. The basic circumstance giving rise to market failure in the innovation sector is the positive externality associated with information production. Special tax rules do not correct the externality; they merely compensate for it through other mechanisms that themselves create deadweight loss. In place of special tax rules that promote innovation, policy-makers should adopt rules that counteract tax-induced distortions in the innovation sector that are disproportionately large. Among these distortions is excess risk-taking, a phenomenon attributable to the interaction between the lognormal nature of returns to risk-bearing and incomplete loss offsets. The absence of full loss offsets can be expected to induce some investors to take on excess risk, thereby increasing deadweight loss and creating negative externalities.

TABLE OF CONTENTS

I. INTRODUCTION 1044
II. REGULATORY APPROACHES TO SOLVING INNOVATION PROBLEMS 1048
  A. *Standard Tax Approaches* 1048
    1. *Equity Considerations* 1049
    2. *Efficiency Considerations* 1054
  B. *"User Pays" Regimes* 1057
III. TAXATION IN AN INNOVATING ECONOMY 1059
  A. *Simple Economic Model of Innovation* 1060
    1. *Nature of Returns in the Innovation Sector* 1061
    2. *Effect on Noninnovation Sector* 1066
  B. *Introduction of Taxes* 1066
    1. *Ideal Taxation* 1067

* Professor of Law, University of Colorado. Thanks to Erik Gerding, Jon Hamilton, Susan Morse, Alex Raskolnikov, Diane Ring, Dan Sokol, Sloan Speck, Harry Surden and participants at various colloquia. The University of Colorado provided research support. I remain responsible for any errors.

2. *Pure Income Taxation* 1069
a. Labor Supply 1070
b. Capital Investment 1071
3. *Real Income Taxation* 1073
4. *Incentive Effects* 1075
a. Shift out of the Innovation Sector 1076
b. Increased Diversification 1076
c. Shift to Higher Expected-Payoff Activities 1078
C. *Pairing Full Loss Offsets with Higher Rates on Innovation* 1083
IV. FULL LOSS OFFSETS: *EX ANTE* AND *EX POST* 1085
V. CONCLUSION 1089

I. INTRODUCTION

The taxation of innovation generally, and of intellectual property specifically, has received a great deal of scholarly attention over the last twenty years.[1] The outpouring of scholarship is related both to the transformation of our industrialized economy into an information economy, and to the distinctive features of information production, particularly the belief that it suffers from endemic market failure.[2] Market failure in the innovation setting takes the form of underproduction due to innovators' inability to capture all of the benefits of their innovative efforts. Others capture some of these benefits because information—the chief product of most innovative activity—is a good that is often, at least partly, nonrival and nonexcludible in consumption.[3] Nonrivalry means that once the good is supplied at any level, its marginal cost drops to zero; it costs nothing to provide it to the next consumer. Nonexcludability means that once it is furnished to anyone, it is furnished to everyone. When goods are nonrival and nonexcludible, many of those who benefit from them do not pay for them, which means that producers of such goods are underpaid. Consequently, policy-makers worry that if legal rules encouraging information production are not in place, too little information will be produced.

1. *See, e.g.*, Andrew Blair-Stanek, *Intellectual Property Law Solutions to Tax Avoidance*, 62 UCLA L. REV. 2, 4 (2015); Michael J. Graetz & Rachael Doud, *Technological Innovation, International Competition, and the Challenges of International Income Taxation*, 113 COLUM. L. REV. 347, 351 (2013); Richard S. Markovits, *On the Economic Efficiency of Using Law to Increase Research and Development: A Critique of Various Tax, Antitrust, Intellectual Property, and Tort Law Rules and Policy Proposals*, 39 HARV. J. LEGIS. 63, 68 (2002). A recent, more general search in Westlaw's Journals and Law Reviews database for articles having "tax" or "taxation" in the title and "innovation or "intellectual property" in the body of the text returned 562 results.

2. *See, e.g.*, John F. Duffy, *Rethinking the Prospect Theory of Patents*, 71 U. CHI. L. REV. 439, 440–43 (2004); Noam Noked, *Integrated Tax Policy Approach to Designing Research and Development Tax Benefits*, 34 VA. TAX REV. 109, 114–15 (2014).

3. Benjamin Russo, *Innovation and the Long-Run Elasticity of Total Taxable Income*, 75 S. ECON. J. 798, 799–800 (2009).

Most of the tax scholarship in the area begins from the seemingly sensible proposition that policy-makers ought to use the tax law as one among several policy instruments available to ameliorate the problem of information underproduction.[4] Adopting or expanding special tax rules that in some way favor innovation will result in more information production, thereby mitigating the market failure.[5] Among the rules considered have been reduced tax rates on income from innovation, enhanced tax deductions for the costs of producing innovation, and various sorts of tax credits for investment in innovative activity.[6] Much of the literature is devoted to assessing the merits of these regimes relative to each other, while another branch compares them to various nontax regimes having the same general objective, such as the patent system and the laws of trademark and copyright.[7]

Legislative developments and proposals in the tax area have run parallel with the scholarship.[8] Provisions currently in effect designed to further innovation include the research and experimental expenditure deduction,[9] the "R&D tax credit,"[10] and capital-asset treatment of gains realized by individuals on disposition of self-created patents.[11] Additional legislative proposals include reduced tax rates for income from innovative activity (so-called "innovation boxes"), and other favorable rules, such as "super deductions."[12]

This Article argues for a different approach to the taxation of innovation. Tax rules should not be designed specifically to further innovation but rather should be adjusted to remove preexisting tax distortions operating in the sector that excessively retard innovation—"excessively," that is, relative to the tax distortions in the noninnovation sector. Three principles support the idea that tax rules ought not be specifically tailored to promote innovation. The first is that the goal of fairness in the taxation of innovative activity, often cited as a reason for special tax rules,[13] is not properly a focus of sector-oriented reform. A number of proposals weigh fairness, efficiency, and other considerations in attempt-

4. *See generally supra* text accompanying note 1.

5. *See, e.g.*, Daniel J. Hemel & Lisa Larrimore Ouellette, *Beyond the Patents-Prizes Debate*, 92 TEX. L. REV. 303, 307 (2013); Xuan-Thao Nguyen & Jeffrey A. Maine, *Equity and Efficiency in Intellectual Property Taxation*, 76 BROOK. L. REV. 1, 3 (2010) (focusing on fairness and efficiency of tax rules for intellectual property).

6. *See* Graetz & Doud, *supra* note 1, at 353 (identifying instruments).

7. Hemel & Ouellette, *supra* note 5, at 307.

8. *See, e.g.*, U.S. S. COMM. ON FINANCE, THE BUSINESS INCOME TAX BIPARTISAN TAX WORKING GROUP REPORT 49–51 (2015), http://www.finance.senate.gov/download/?id=B4AEDDC8-9E94-4380-9AF4-9388953FB347. (considering the possible benefits of a patent/innovation box regime). For a general description and assessment of tax inducements to innovation, see Graetz & Doud, *supra* note 1, at 351–52.

9. I.R.C. § 174 (2012). Unless otherwise stated, all section citations are to the U.S. Internal Revenue Code of 1986, as amended (26 U.S.C.).

10. *Id.* § 41.

11. *Id.* § 1235.

12. *See generally* Graetz & Doud, *supra* note 1.

13. *See, e.g.*, Nguyen & Maine, *supra* note 5, at 3.

ing to craft an appropriate tax regime for innovation.[14] Fairness, however, is ideally a property of the system as a whole, not of any individual sector and almost certainly not of an arbitrarily defined economic sector. Requiring fairness within a particular economic sector simply limits reform options with no apparent payoff, given that other parts of the system are likely to remain unfair (and possibly to become more so) following the adoption of a targeted fairness reform. Indeed, an extensive literature has persuasively argued that it is appropriate to address and evaluate fairness globally rather than sector-by-sector, and to address it using government transfer payments based on each individual's overall welfare rather than by targeted subsidies, however delivered.[15] Moreover, even if a robust system of transfer payments is unavailable for political reasons, fairness would be better addressed using distributional measures that take account of all of an individual's economic resources, not just those located in one sector. By separating the efficiency question from the fairness question, a better menu of reform options is available.

The second principle is the recognition that correcting an externality does not necessarily lead to an improvement if the method of correction requires the introduction of distortions into other parts of the system. A proposal that increases one set of distortions to reduce another requires us to know whether we have lost more by robbing Peter than we have gained by paying Paul. Furthermore, even if we have not lost more, the question arises whether Peter must be robbed in the first place. A sectorial approach to regulating innovative activity of the kind advocated here avoids this problem to a great extent because it focuses first on correcting tax distortions already operative in the innovation sector. Because a correction of these distortions can be welfare-enhancing by itself, not only may it become unnecessary to rob Peter to pay Paul, but it may be possible to pay them both.

The third principle is, in a sense, a corollary of the second: special tax rules geared toward stimulating innovation suffer from a structural handicap relative to other possible regimes because special tax rules do not address the basic circumstance leading to market failure for innovation. As explained above, that circumstance is the problem of spillover effects, or free-riding by beneficiaries, of innovation.[16] If market failure in the innovation, sector is due to spillover effects of innovation, the solution is to make the third parties enjoying the spillover bear the marginal costs that give rise to the benefits they enjoy, not to stimulate investment with resources from others.[17] Existing nontax regimes in the innovation

14. *Id.* at 27; Hemel & Ouellette, *supra* note 5, at 303.

15. *See, e.g.*, JOHN RAWLS, A THEORY OF JUSTICE 3 (1999) ("Justice is the first virtue of social institutions . . . ."); *see also* LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE (2002) (arguing for the use of a global welfarist approach to evaluate social institutions).

16. *See* Hemel & Ouellette, *supra* note 5, at 330, 368.

17. *See* Donald B. Marron & Adele C. Morris, *How Should Governments Use Revenues from Corrective Taxes?*, URBAN INST. & BROOKINGS INST. TAX POLICY CTR. (Jan. 28, 2016), http://www.taxpolicycenter.org/publications/how-should-governments-use-revenue-corrective-taxes/full.

sector, however imperfect, are structurally designed to achieve this goal, whereas generalized income-tax rules cannot. Patent, copyright, and trademark holders collect elevated profits from the users of the protected property, but tax subsidies transfer funds from general revenues to innovators, not specifically from the beneficiaries of innovation.[18] Any resulting increase in innovation investment due to special tax rules is consequently offset by increased tax distortions from higher taxes in other sectors.[19]

From this perspective, the tax question on innovation is not how to design rules that encourage it, but, as suggested, how to remove tax distortions that excessively discourage it. The primary driver of additional distortion in the innovation sector—and indeed for all investments subject to uncertain returns—is the nondeductibility of overall losses (usually termed the absence of "full loss offsets").[20] As contrasted with the fixed-return sector, in which net losses by definition do not arise as long as the risk-free rate is nonnegative, an endemic feature of the innovation sector is the possibility that the taxpayer will have negative income. Ventures often fail, and, even if they do not fail, they commonly experience losses before they enjoy gains. When taxpayers must factor in the possibility of nondeductible losses from an activity against taxable gains, the expected, or *ex ante*, tax cost of the activity rises,[21] leading, in turn, to at least three first-order substitution effects. The first and most obvious is an exodus of factors of production from the innovation sector as taxpayers seek a more favorable expected after-tax return in the noninnovation sector.[22] A second effect is a tendency toward diversification in the innovation sector, as affected parties attempt to ensure that they have some positive income against which they may offset otherwise unusable losses.[23] The mechanism is likely to be pursuit of tax-motivated business combinations.[24] A final effect is, perhaps counterintuitively, an excess of risk-taking in the sector as firms seek to reduce their expected tax rate by taking bets with larger expected gains on projects should they succeed, relative to expected losses should they fail; this effect results from the lognormal nature of returns to risk-taking.[25] There is substantial evidence of the presence of the first two effects and that they are significant; evi-

18. Hemel & Ouellette, *supra* note 5, at 309.

19. The deadweight loss or "excess burden" of a tax is the social cost, in the form of otherwise-efficient forgone transactions, resulting from the tax. *See infra* Subsection III.B.2.

20. Mark Campisano & Roberta Romano, *Recouping Losses: The Case for Full Loss Offsets*, 76 NW. U. L. REV. 709, 729 (1981).

21. Jacob Nussim & Avraham Tabbach, *Tax Loss Mechanisms*, 81 U. CHI. L. REV. 1509, 1558 (2014).

22. *See* Campisano & Romano, *supra* note 20, at 722–30, for a general discussion of the problem.

23. *Id.* at 721–22.

24. *Id.* at 710.

25. Under a lognormal distribution, losses relative to the median return are less, on average, than are gains. *See infra* Subsection III.A.1. To my knowledge, the point that loss limitations can lead to an increase in risk-taking has never been made in either the economics literature or the legal literature.

dence also supports the existence of the third effect, although the extent to which it is due to tax distortions or to other circumstances is less clear.[26]

This Article addresses these issues as follows. Part II reviews literature on taxation of innovation—offering some criticisms in light of the framework developed in Part III. Part III presents a simple economic framework for analyzing the nature of innovation, examines the effects of taxation on innovation, and then discusses tax policy options. Part IV focuses specifically on the question of whether *ex ante* or *ex post* incentives for innovation are superior, arguing that the *ex ante* view developed in Part III finds further support in recent property theory.

II. REGULATORY APPROACHES TO SOLVING INNOVATION PROBLEMS

The animating principle of most of the tax scholarship on innovation is that tax rules ought to address the market failure endemic to innovation.[27] The market failure takes the form of a positive externality, or spillover; it results from the fact that once information is produced, it can be appropriated at little or no cost by third parties. Those parties value the information and would pay for it if they were compelled, but because they are not compelled, information producers end up undercompensated. The overall effect is underinvestment in innovation.

A. Standard Tax Approaches

Standard approaches to the problem of underinvestment in the innovation sector evaluate a range of available instruments, seeking to identify those that have the greatest "bang for the buck" in terms of stimulating investment and, often, in terms of other criteria, such as promoting horizontal and vertical equity.[28] This Section explores some of the difficulties with tax rules designed to promote these ends, arguing that none of them optimally address the basic problem of market failure in the innovation sector because either they do not seek to properly internalize costs, or they improperly focus on equity concerns that policymakers would better address on a more global scale.

26. *See, e.g.*, William M. Gentry & R. Glenn Hubbard, *Tax Policy and Entrepreneurial Entry*, 90 AM. ECON. REV. 283, 283 (2000) (demonstrating dampening effect of progressive rates on entry into entrepreneurial ventures); Campisano & Romano, *supra* note 20, at 730–31 (noting that tax-created incentive of businesses to merge or combine before losses are sustained in order to increase the likelihood that losses of one business will be currently deductible).

27. The problem of market failure in the innovation sector is effectively canonical. *See* Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609–10 (1962).

28. *See supra* note 1.

1. Equity Considerations

Two recent articles, one by Xuan-Thao Nguyen and Jeffrey Maine, and one by Daniel Hemel and Lisa Ouellette, are characteristic of much of the tax scholarship that focuses on balancing equity and efficiency in the taxation of innovative activity. At the outset of their analysis, Nguyen and Maine state: "Ideally, the current intellectual property tax system should embrace the principles of fairness and efficiency,"[29] and they proceed to analyze the system in those terms, recognizing that the principles may be in tension.[30] Hemel and Ouellette proceed in similar fashion, except they broaden the analysis to compare the relative merits of tax and nontax instruments, such as patents, prizes, and grants.[31] Like Nguyen and Maine, their focus is on both efficiency and fairness considerations, though they are more concerned with vertical than horizontal equity.[32]

A focus on the asserted horizontal inequities of various tax rules with regard to innovation seems misplaced. As an example, consider certain differences in the tax treatment of IP acquisition costs that Nguyen and Maine identify.[33] They note that an inventor who incurs such costs in connection with the development of a patent may deduct them if the inventor plans to market the patent in the inventor's own trade or business, but must capitalize them if the inventor plans to license the patent to third parties.[34] Under a capitalization regime, deductions for the cost of the invention must be taken over time rather than upfront,[35] thereby increasing the real after-tax cost of invention.[36] Nguyen and Maine assert that the disparate treatment is inequitable on the basis that otherwise identically situated inventors experience different tax treatments depending on the way each decides it best to generate income.[37] The licensor faces a lower after-tax return than does the inventor who directly uses the patent in trade or business.

Focusing solely on the equity question, one can imagine at least three reasons why the objection seems misplaced, and these reasons apply more generally to localized concerns about equity. The first is that market effects should compensate a licensor for most or all of the tax disadvantage that Nguyen and Maine identify. A potential innovator contemplating development of a patent for license will factor in the add-

29. Nguyen & Maine, *supra* note 5, at 3.
30. *Id.* at 6.
31. Hemel & Ouellette, *supra* note 5, at 307 (stating that they present a new framework to answer the questions of when society should use intellectual-property instruments and when it should use tax instruments to stimulate intellectual-property innovation).
32. *See id.* at 303.
33. Nguyen & Maine, *supra* note 5, at 15–19.
34. *Id.* at 16.
35. *See* I.R.C. § 262 (capitalization requirement); *id.* § 167 (depreciation generally).
36. Nguyen & Maine, *supra* note 5, at 2–3 (discussing the shortcomings of traditional taxation principles as applied to intellectual property).
37. *Id.* at 14.

ed tax cost.[38] The effect will be to reduce the number of licensors, thereby increasing the return to licensors and driving down the price of inputs.[39] In other words, as long as markets are reasonably efficient, compensating market effects eliminates the inequity. Of course, there are plenty of adverse efficiency effects, but the focus here is on equity.

Secondly, even if markets are not reasonably efficient, so that compensating market effects do not arise, the inequity seems to be a transition concern rather than a concern about the IP sector in a steady state. As a general matter, no one is compelled to be a licensor rather than a direct marketer of patented IP. Consequently, disadvantaged persons are those who already have invested in a patent-license decision at the time the adverse rules come into effect, but who lacked a basis at the time of their investment to know that the rule would change; others simply do something else. The treatment of economic actors disadvantaged because of prior costly decisions is the subject of the large literature on the economic effects of legal transitions.[40] The general focus of the literature is on the questions of whether and to what extent private actors ought to be compensated for the costs of legal change.[41] But, however those issues are resolved, they are not fundamentally about the inequity of steady-state rules; rather, they are about combating the inequities and inefficiencies associated with reasonable reliance on prior states of affairs when rules change.[42]

Finally, the fact of disparate rules in the IP sector (to use Nguyen and Maine's case) does not seem relevant to determining whether there is inequity in the system as a whole unless the only way to redress global inequity is sector by sector. Setting that hypothesis aside for the moment, the question becomes whether local "inequity" is troubling on its own terms (and, again, keeping the analysis separate from that of allocative inefficiency resulting from the same rules). Even if fairness or justice is generally considered to be a property of particular institutions, rather than of society generally, it is not obvious why inequity within a particular, and to a large extent arbitrarily defined, economic sector ought to matter, assuming the purpose of activity in the sector is to promote the most just distribution of social wealth overall. If the purpose includes other values, such as ensuring a satisfying mode of self-expression, then perhaps sectorial equity counts. That, however, seems to raise a question of political theory or philosophy beyond the scope of a discussion of tax rules. If the purpose is to promote the most just distribution, then sectorial inequity that is productive of more social wealth is not merely ac-

38. *See id.* at 16.

39. This is because the higher tax costs will reduce the number of licensors.

40. *See, e.g.*, DANIEL N. SHAVIRO, WHEN RULES CHANGE: AN ECONOMIC AND POLITICAL ANALYSIS OF TRANSITION RELIEF AND RETROACTIVITY (2000); Louis Kaplow, *An Economic Analysis of Legal Transitions*, 99 HARV. L. REV. 509 (1986).

41. *See, e.g.*, Kaplow, *supra* note 40, at 511.

42. *Id.* at 513.

ceptable; it is preferable to sectorial "fairness" that results in less total wealth, assuming it is possible to redress the distributional effects of the efficient rule through redistribution effectuated by other means, such as government transfer payments.[43] The only case in which that proposition would not hold would be where there is an inadequate global framework to ensure a just distribution of social wealth, so that system-wide equity had to be built up from equity within individual parts of the system.

Hemel and Ouellette, perhaps agreeing with some of these observations, do not focus as much on horizontal equity as on vertical equity—arguably the more relevant equity concern as a general matter[44]—and they also focus more heavily on systemic efficiency considerations than do Nguyen and Maine. Consequently, they adopt a somewhat different approach from Nguyen and Maine to developing tax rules for innovation. Hemel and Ouellette argue that the government has the capacity to promote innovation along three axes: the amount of the benefit, its timing, and who pays for it.[45] In any individual case, optimally promoting innovation involves deciding from among the axes and along any one of them, with equity and efficiency considerations potentially operative along each dimension.[46] They observe, for example, that tax benefits such as accelerated depreciation or tax credits represent transfers from society as a whole to innovators, whereas patent rights reflect a "user pays" principle ("who pays").[47] The former spreads costs but also compels nonusers to pay, whereas the latter may permit inventors to capture monopoly rents that prevent some people from purchasing who are willing or able to pay the marginal cost but not the monopoly cost.[48] Hemel and Ouellette also identify distributional and efficiency differences between *ex ante* incentive mechanisms (refundable tax credits, grants) and *ex post* mechanisms (prizes, patents, reduced tax rates).[49] The former are available regardless of the success of the innovation, whereas the latter apply only at the back

43. *See* Hemel & Ouelette, *supra* note 5, at 346–52.

44. *See* Louis Kaplow, *Horizontal Equity: Measures in Search of a Principle*, 42 NAT'L TAX J. 139 (1989) (arguing that horizontal equity is a normatively unfounded ethical principle).

45. Hemel & Ouellette, *supra* note 5, at 308–09.

46. *Id.*

47. *Id.* at 308. Hemel and Ouellette state that the user-pays principle generates more deadweight loss than does an R&D tax credit with broad-based financing. *See id.* ("From an efficiency perspective, spreading the costs of R&D across a broad base [rather than through patents] will reduce the corresponding deadweight loss."). This claim seems to be incorrect. The user-pays principle properly internalizes costs if the user is the one who benefits from the innovation and is accurately charged the costs associated with the benefit enjoyed. In that case, there is no deadweight loss. This is simply a restatement of the first fundamental theorem of welfare economics. *See* Kenneth J. Arrow & Gérard Debreu, *Existence of an Equilibrium for a Competitive Economy*, 22 ECONOMETRICA 265, 286 (1954). The general rule that a broad-based tax reduces deadweight loss applies in the case of governmental provision of public goods and to finance redistribution, not goods for which it is possible to identify positive marginal costs (to inputs) and that are to be privately supplied. HARVEY S. ROSEN & TED GAYER, PUBLIC FINANCE 329–51 (9th ed. 2010).

48. Hemel & Ouellette, *supra* note 5, at 314 (discussing deadweight losses associated with monopoly and taxation).

49. *Id.* at 308.

end, once the taxpayer has successfully innovated.[50] The efficiency effects differ in that, under an *ex ante* regime, a greater incentive for more speculative innovation or for projects tangential to the intended innovation arises; differences in equity effects include a transfer from society as a whole to all innovators in the case of *ex ante* incentives, or just to successful ones in the case of *ex post*.[51] In Hemel and Ouellette's view, the optimal incentive in any particular case consequently depends on a variety of case-specific considerations.[52] These may include whether distributional considerations militate against patent protection (where, for example, the patent is for life-saving technology and should be made widely available), and whether factors such as investor risk aversion counsel in favor of smaller but more certain benefits or larger ones contingent on investor success.[53] For all proposals, however, they consider that there is an efficiency tradeoff. For example, they state that any tax favoring innovation will create some deadweight loss,[54] while even a patent system will, in their view, result in suboptimal provision of the patented good because of monopoly pricing.[55]

Like Nguyen and Maine, Hemel and Ouellette suppose that the optimal approach to taxing innovation is to assess the tradeoffs between or among various policy instruments in light of the goals of promoting efficiency and equity;[56] though Hemel and Ouellette properly, in my view, evaluate the equity question in systemic terms, not in terms of the similarity of treatment of parties similarly situated in the innovation sector.[57] Again, however, the question remains why they begin with proposals that require efficiency tradeoffs, and it remains unclear why they roll equity considerations into their analysis.

Consider, for example, Hemel and Ouellette's discussion regarding the choice of whether to use patents or prizes to promote innovation.[58] Patents operate on a "user pays" principle, whereas prizes typically are funded out of general tax revenues.[59] (Prizes need not be so funded, but to the extent they are funded out of measures that look to users of the rewarded innovation for payment, such as a tax on the rewarded good, they resemble patents.[60]) If the innovation is one that ought to be available to all who need it, such as an AIDS vaccine, then a patent is less desirable, in their view, than a prize because under a patent, the innovator recoups its cost through monopoly pricing; as a consequence, some users

50. *Id.*
51. *Id.*
52. *Id.* at 376.
53. Hemel & Ouellette, *supra* note 5, at 326–67.
54. *Id.* at 314.
55. *Id.* at 349–50. The next subsection questions whether "monopoly" profits in the innovation sector are in fact inefficient.
56. *Id.* at 309.
57. *See generally id.*
58. *Id.* at 376–77.
59. *Id.*
60. *Id.* at 312.

may not be able to afford the drug.[61] A prize funded out of general tax revenues will therefore be more effective than a patent because the technology becomes immediately available to numerous suppliers who compete with the inventor on price.[62] By contrast, in the case of innovations that more closely resemble simple consumer goods, such as a baldness cure or a better television set, the opposite rule is preferable because there is no apparent reason why anyone ought to be considered entitled to the good.[63] Rather, from both an efficiency and equity standpoint, the "user pays" principle seems appropriate.[64]

As discussed above, equity is fundamentally a question of how societal resources are distributed.[65] It raises a different set of concerns than does efficiency—concerns that are more effectively addressed with different measures. An approach that combines the two considerations is likely to end up needlessly reducing total social welfare because of the associated efficiency costs, while, in many cases, failing to achieve the desired equity objective (either by falling short of the target or by overcompensating for distributional problems). Conversely, a generalized approach that separates the effort to reach an equitable distribution from measures designed to promote efficiency is likely to avoid unnecessary welfare losses without any increase in distributional inequity.

As an illustration, consider the difference between prizes for innovation financed with general tax revenues and patents.[66] Regardless of whether monopoly profits associated with patents create deadweight loss (addressed in the next Subsection),[67] a prize funded by a general tax increases deadweight loss because of the substitution effects toward leisure and current consumption that comes with any income tax.[68] Further, as the tax increases, the loss increases approximately as the square of the tax.[69] The asserted equity benefit of the prize is that purchasers will not have to pay the monopoly cost that patent rights provide.[70] The benefit would be important in the case, for example, of a medically necessary

61. *Id.* at 313.

62. *Id.* at 308.

63. *Id.* at 350. Entitlement to the good should not be confused with entitlement to an overall basket of resources that might guarantee individuals generally with resources that would enable them to purchase certain consumer goods. This type of consideration would be the focus of more general arguments about the proper allocation of goods in society.

64. The principle of "user pays" may be considered a statement of the first fundamental theorem of welfare economics. *See* Allan M. Feldman, *Welfare Economics*, *in* 4 THE NEW PALGRAVE DICTIONARY OF ECONOMICS 889 (John Eatwell et al. eds., 1987).

65. *See supra* Part I.

66. For a discussion of the relative merits of prize and patent regimes, see, for example, Steven Shavell & Tanguy Van Ypersele, *Rewards Versus Intellectual Property Rights*, 44 J.L. & ECON. 525 (2001).

67. As I argue in Section II.B below, a properly structured patent regime is actually more efficient than a prize or government subsidy regime.

68. *See infra* Subsection II.A.2.

69. ROSEN & GAYER, *supra* note 47, at 338. *See infra* Subsection II.A.2 for an explanation of the squaring rule.

70. Hemel & Ouellette, *supra* note 5, at 381.

drug, though less so, or perhaps entirely absent, in the case of a normal consumer good.

But note that there is no obvious reason why the same mechanism that ensures the development of the drug (a prize) ought also to ensure access. The relationship between the equity benefit delivered by a prize regime and the equity concern the prize system is supposed to address is essentially arbitrary. Under the prize system, drug prices are asserted to be "correct" in the sense that monopoly profits are stripped out of the market (ignoring the deadweight loss due to financing the prize).[71] But, even if the price is correct (a debatable point), there is no guarantee, or even reason to assume, that the drug is appropriately available. Typically, merely stripping out monopoly profits will overshoot or (more likely) undershoot the objective of making a given drug affordable for all who need it.[72] This problem would arise, in particular, where the noninformation costs of providing many drugs may be considerable.[73]

These considerations suggest that it would make more sense for policy-makers to adopt a system of social insurance that generally provides for medical care to those who cannot afford it than to use a distortion-inducing mechanism (taxes to fund prizes) to finance innovation to ensure access. The considerable economies of scale associated with a general system of social insurance are likely to reduce the overall cost of the insurance, even if funded by general tax revenues.[74] If Congress were to deal with the access problem through generalized insurance it could consider solely the optimal efficiency rule for the innovation. That rule is some version of user pays, not a general income tax rule. Among the candidates are patents, where users are charged directly, and prizes and grants that are funded through excise taxes on the goods developed with the prize or grant.

2. *Efficiency Considerations*

Difficulties also arise in connection with efforts to improve efficiency through the use of targeted tax rules. As explained previously, the characteristic problem for information production is the market failure associated with the nonexcludability of information.[75] Once information is produced, third parties may be able to appropriate it at little or no cost, meaning that innovators are undercompensated.[76] As a consequence, they leave the innovation sector, resulting in underproduction.[77]

71. *Id.*
72. *Id.* at 349.
73. *Id.*
74. *Id.* at 380
75. *See supra* Part I.
76. Arrow, *supra* note 27, at 614–15.
77. *Id.* at 617.

The standard approach of policy-makers and scholars in the tax area to address the problems is to consider various targeted tax rules that promote information production.[78] These rules may reduce the after-tax costs of innovation or increase its after-tax benefits. As examples of the former, at various points Congress has provided for a research and experimentation credit[79] (recently made permanent),[80] and for a deduction for certain qualifying research and experimental expenditures.[81] Examples of the latter include patent boxes[82] (proposed in the U.S. but not yet adopted),[83] and capital-asset treatment for certain forms of intellectual property.[84] In both cases, innovation becomes more attractive through a government subsidy to innovators.

Although it appears that regimes of this type do tend to promote innovative activity to varying degrees,[85] the question remains whether they are, on balance, helpful given their necessarily distortionary character in other sectors. Because they distort, it becomes necessary to weigh the efficiency losses they create against the efficiency gains they supply. The distortion results from the fact that the use of generalized taxes to fund specific industries creates a cross-subsidy, in that the beneficiaries of the subsidy—users of information—do not pay for it.[86] Instead, the cost is spread throughout the system in the form of somewhat higher tax rates.[87]

The mechanism by which social costs arise is comparable. Any real tax used to fund a generalized benefit will have a substitution effect. Because the taxed activity becomes more expensive, taxpayers at the margin substitute another activity for the taxed one once the after-tax value of the taxed activity has fallen below the value of the substitute.[88] Assuming that the pretax allocation of resources was efficient, it follows that the

78. *See supra* note 1.

79. I.R.C. § 41 (2012).

80. Pub. L. No. 114–113, § 121(a)(1) (2015).

81. I.R.C. § 174.

82. According to one study, as of 2015, patent box regimes were in effect in thirteen European countries and in China. Sebastien Bradley et al., *Cross-Country Evidence on the Preliminary Effects of Patent Box Regimes on Patent Activity and Ownership* 1 (Tuck Sch. of Bus., Working Paper No. 2681433) (October 2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2681433.

83. JOINT ECONOMIC COMMITTEE, PATENT BOXES: A BRIEF HISTORY, RECENT DEVELOPMENTS, AND NECESSARY CONSIDERATIONS (Mar. 10, 2016), https://www.jec.senate.gov/public/_cache/files/02a2a18a-1e08-42ce-8c14-72b6138b54dd/031016-patent-boxes.pdf.

84. I.R.C. § 1235.

85. *See* Bradley et al., *supra* note 82, at 15 (finding a 3% increase in patent applications for every 1% reduction in the tax rate on patent income); Graetz & Doud, *supra* note 1, at 347 (arguing that R&D incentives and not patent box regimes are effective for the related purpose of stimulating productive activity in the jurisdiction in which enacted).

86. *See* William P. Kratzke, *Tax Subsidies, Third-Party-Payments, and Cross-Subsidization: America's Distorted Health Care Markets*, 40 U. MEM. L. REV. 279, 349–50 (2009) (discussing problems with cross-subsidization).

87. Stanley S. Surrey, *Tax Incentives As A Device for Implementing Government Policy: A Comparison with Direct Government Expenditures*, 83 HARV. L. REV. 705, 725–26 (1970) (discussing the higher tax rates caused by the use of tax incentives).

88. *See* ROSEN & GAYER, *supra* note 47, at 336, for an explication of this standard point.

after-tax allocation is less efficient, with the loss of welfare measured by the extent of surplus forgone from the substitution.[89] This principle applies even in the case of an income tax, which, although it reaches all goods and services, does not tax leisure and does not further burden current consumption.[90] If leisure remains untaxed, taxpayers will consume more of it as the cost of acquiring income rises from taxation. Similarly, because current consumption remains more favorably taxed than future consumption under the tax, one can expect taxpayers to substitute present consumption for additional savings.[91]

Moreover, the distortionary effect of these substitutions is exponential because an increase in tax has effects on two margins: it increases both the number of substitutions that taxpayers make and the amount of economic surplus forgone with each additional substitution.[92] As indicated above, the exponentially rising nature of the welfare loss is sometimes referred to as the "squaring principle" of deadweight loss.[93] It means, for example, that if a tax at rate t generates $\$D$ of deadweight loss, a tax of $2t$ will generate approximately $\$4D$ of deadweight loss.

The implication of the analysis for targeted tax rules is that they come at a cost in the form of efficiency losses system-wide. These costs, when combined with administrative, compliance, and, perhaps, other costs,[94] may in some cases exceed efficiency gains from the rule.[95] Even if they do not exceed gains, the net benefit may be small, inviting the question of whether the opportunity cost of the rule exceeds the benefit. This would occur where better alternative methods of addressing innovation market failure are available. As a general matter, better methods would associate payment for the benefits that innovation supplies with the beneficiaries. In that case, their benefits as a group are aligned with the costs they impose, while those who do not enjoy benefits are not required to bear any costs. In other words, to the extent benefits of information production are not universally enjoyed, they resemble market goods and should be financed similarly to the mechanism for financing market goods. This is a simple user-pays regime, discussed below.

89. *Id.* at 337–38.

90. *See infra* Subsection III.B.2.

91. Specifically, faced with the choice of consuming or investing $\$X$ of after-tax income, the taxpayer faces no additional tax on the former but tax on returns to investing on the latter.

92. David A. Weisbach, *Line-Drawing, Doctrine, and Efficiency in the Tax Law*, 84 CORNELL L. REV. 1627, 1656 (1999) [hereinafter Weisbach, *Line-Drawing*] (discussing deadweight loss and efficiency generally).

93. *Id.*

94. Compliance and administrative costs for many targeted tax benefits are notoriously high. *See, e.g.*, Richard Krever, *Analysing Implicit Tax Expenditures,* 35 MELBOURNE U. L. REV. 426, 428 (2011) (noting the high compliance costs associated with poorly designed tax expenditures).

95. When cross-border behavioral adjustments are factored in, the efficiency benefits may be even further reduced. *See* Graetz & Doud, *supra* note 1, at 410.

B. "User Pays" Regimes

In the preceding discussion, I suggested that "user pays" regimes, in particular the substantive laws of patent, copyright, and trademark, are superior to tax regimes as instruments to control the externalities associated with innovation. The suggestion is contrary to most of the literature, which generally takes the view that monopoly profits (secured by patent or other intellectual property (IP) rights) represent deadweight loss, whereas a system that prices innovation at marginal cost does not.[96] The deadweight loss is claimed to arise because monopoly pricing artificially limits demand.[97] In opposition to this claim, I suggest below that monopoly pricing for innovations can (not "does" or "must") reflect an adjustment to the payment stream, not to total demand. Here I sketch out the argument in favor of IP regimes and associated monopoly profits in a suggestive way. The purpose is to demonstrate that the ideal forms of such regimes actually produce no deadweight loss, and, therefore, that other things equal, properly crafted rules in these areas will be superior to generally applicable tax rules in efficiency terms. The motivating intuition is that what are commonly termed "monopoly profits" for returns to innovation in principle may reflect only an adjustment to the timing and source of the payment stream that innovators receive, not the transfer of economic rents to them, assuming the regime gets the cost of providing the innovation correct.[98] To be clear, the claim is not that policy-makers should not take measures to reduce the inequities associated with monopoly pricing for IP, or that monopoly pricing is efficient as a general matter; rather, it is that monopoly pricing may be an effective means of internalizing the spillover effects of innovation. The implication is that policy-makers ought to pursue distributional objectives such as ensuring access to life-saving drugs through mechanisms unrelated to the mechanism by which innovators get paid.

A feature of innovation that tends to obscure the point is that returns to information production generally, although not always, are uncertain or risky and for that reason may not correspond to the actual cost of producing the information in individual cases. Consider that a successful inventor is likely to reap high returns, while an unsuccessful inventor reaps few or none. For the latter, losses are limited to costs, but for the former, returns may vastly exceed costs. This means that some of the return to innovation reflects a simple risk premium. The risk premium, however, is not a part of the externality that is intrinsic to innovation, or

96. *See, e.g.*, Shavell & Van Ypersele, *supra* note 66, at 529–30 (arguing that monopoly pricing of innovations represents deadweight loss); Hemel & Ouellette, *supra* note 5, at 308 (arguing that prizes for innovation are superior to patents because prizes spread the cost of the innovation across society, resulting in less deadweight loss than patents due to the squaring rule).

97. *See* Shavell & Van Ypersele, *supra* note 66, at 529–30.

98. The notion that returns to innovation generally take the form of inefficient monopoly rents is widespread in the literature. *See, e.g.*, Michael Abramowicz, *Perfecting Patent Prizes*, 56 VAND. L. REV. 115, 130 (2003) (characterizing monopoly pricing as "inherent[ly]" producing deadweight loss).

at least to the production of information. Rather, the externality derives from the circumstance that innovation is a public or quasi-public good, as evidenced by the fact that the externality can arise without the risk and the risk can arise without the externality. As an example of the former, one can imagine certain information goods, such as data about the temperature of ocean water at various depths, that might be relevant to a given innovation, but whose cost to acquire and whose payoff from having acquired are known *ex ante*. As an example of the latter, risky returns are endemic to many, if not most, areas of economic activity: stocks fluctuate in value, interest rates vary, and, more generally, various unpredictable factors (such as weather and consumer tastes) routinely affect returns on investment.[99] Nevertheless, the risk-bearing associated with these phenomena is not due to nonrivalry or nonexclusivity.

It therefore becomes important to the analysis to isolate the public-goods feature of returns to innovation from the risk feature in order to figure out the pricing remedy for the market failure. Only the former gives rise to the need for special legal regimes for innovation that address externalities. Focusing on this feature, it emerges that elevated or "monopoly" returns can be conceptualized simply as one way, among others, of altering the timing of payments to innovators to ensure they receive what is in fact a competitive return.

To isolate the problem that the public-goods character of innovation poses, and that the patent regime (for example) is intended to address, consider the case of innovation without risk. In this stylized setting, a potential innovator ("Innovator") faces the question of whether to engage in innovative project P or to seek returns in the noninnovation sector (which can include other risky activities not requiring innovation). The question becomes how to ensure Innovator faces the right incentives to innovate. Innovator knows that if she invests X years in project P, her investment will result in an innovation that produces some quantity of value, $\$V$, where $\$V$ is the total social value of the innovation, net of costs other than those of the innovative activity itself. We further assume that Innovator faces some annual return, $\$W$, in the noinnovation sector. Ideally, we want Innovator to innovate if $\$V > X*\$W$ (in present value terms). By the same token, we do not want to pay Innovator more than $X*\$W$ to innovate.

The difficulty with standard market mechanisms is that, because the innovation is an information good, there will be a free-rider problem if no special legal regime is in place to compensate Innovator for her innovative efforts (or her contribution of capital to innovation).[100] In the case

99. *See, e.g.*, John R. Brooks II, *Taxation, Risk, and Portfolio Choice: The Treatment of Returns to Risk Under a Normative Income Tax*, 66 TAX L. REV. 255, 264 (2013) (discussing risk and investments generally).

100. *See, e.g.*, David Booton, *The Management of Uncertainty and the Patent Incentive*, 24 I.P.J. 79, 80 (2011) (describing the free-rider problem traditionally associated with the provision of information goods).

of a pure innovation, she will get paid $0 if she innovates and no special legal regime guarantees her return. An ideal patent regime, therefore, would provide a net payment to her equal to X*W (in present value terms). An ideal prize or grant regime would do the same. The differences from Innovator's perspective concern only the manner of payment. Under a patent regime, Innovator can prevent others from manufacturing the product and thereby charge monopoly prices, ideally just until she receives X*W. Under a prize or grant regime (which reduce to the same thing, apart from timing, where there is no risk to innovation), Innovator, in effect, becomes the government's employee or independent contractor, getting paid either before or after she innovates.

Significant welfare differences exist, however, between using the patent regime together with government transfer payments on one hand, and prize or grant regimes on the other, to reach desired distributional outcomes. The use of government subsidies, such as grants or prizes, ties the redistributive relief to the market price. That is, a grant or prize can be used to bring down the consumer's cost of purchasing the innovation to its marginal cost. But setting the price at marginal cost does not necessarily address the distributional concern—there is no systematic relationship between consumer need and market prices. Many consumers who need a product also cannot afford to pay even at cost; others who need it may be able to pay something above a market price, but are unable to pay the monopoly cost. For them, relief should take the form of a price set somewhere between the market and monopoly prices. In short, directly addressing distributional concerns through subsidies for the innovation needlessly ties the government's hands. If, instead, the government uses the patent system (assuming it is able to determine the true cost of the innovation) and separately provides for transfer payments based on overall assessments of individual consumers' needs, a more efficient outcome results.

III. TAXATION IN AN INNOVATING ECONOMY

To this point I have argued that targeted tax rules designed to promote innovation typically suffer from one or both of two structural defects. In the first place, they may be designed to promote equity in the innovation sector, resulting in substantial, unnecessary efficiency losses as compared to regimes that decouple equity from efficiency, addressing the former on a more global basis.[101] Secondly, they may fail to address the problem intrinsic to innovation that necessitates regulatory solutions, which is the externality associated with information production.[102] Consequently, most tax regimes designed to promote innovation needlessly trade off equity and efficiency, and come at the cost of introducing new

101. *See supra* Section II.A.
102. *See supra* Section II.B.

distortions in other sectors, as rates must be raised there to pay for tax benefits for innovation. These considerations suggest that one ought to approach the use of targeted tax rules to promote innovation with caution.

Nevertheless, the perhaps tentative conclusion that targeted tax rules for promoting innovation have only limited utility does not put to rest the question of whether tax reform has a place in addressing problems with information production. Rather, the question remains whether, and to what extent, standard tax instruments affect innovation for good or ill. As I have suggested, there is reason to think that one standard feature of income-tax systems has particularly detrimental effects on innovation (as compared to their effects on noninnovative productive activity) and therefore that sector-oriented tax reform may be in order as a means to promote efficiency in the innovation sector, without loss of efficiency in the noninnovation sector. This feature is the limitation on overall losses.[103]

In order to motivate the analysis, I begin with a simple model of innovation. I then turn to the effects of taxation on innovation, beginning first with ideal taxation and then moving to the actual tax system to understand how loss limitations can dramatically retard innovative activity and how sector-oriented reform could produce substantial welfare improvements.

A. Simple Economic Model of Innovation

It is useful to begin with an intuitive model to understand how innovation functions in a simple economy in which no distortions are present operate. In this setting, perfect competition characterizes the market for all goods and services, and tax revenues are exogenously supplied; no externalities operate to distort returns (including externalities for innovation). Individuals can own capital, provide labor, or both. Individuals differ in consumption tastes but not in risk tolerance, which is initially assumed to be zero. They also do not differ in ability—sometimes termed "wage rate."[104] Because innovation generally requires risk tolerance, there is no innovation sector in the economy.[105] Consequently, the market-clearing price for each good will be equal to the marginal cost to produce it.[106] Marginal cost for a unit of a good includes the cost of all

103. *See* I.R.C. § 1 (2012) (providing for no negative tax on negative income); *id.* § 382 (limitations on trafficking in NOLs); *id.* § 1211 (limitations on capital losses).

104. Kirk J. Stark, *Enslaving the Beachcomber: Some Thoughts on the Liberty Objections to Endowment Taxation*, 18 CAN. J.L. & JURIS. 47, 51 (2005).

105. Technically, innovation can arise without risk, but only if special legal regimes are in place to provide an appropriate return. For present purposes, I assume that innovation is inherently risky. *See infra* Part IV.

106. *Marginal Revenue and Marginal Cost Relationship for Monopoly Production*, BOUNDLESS ECON. (Aug. 8, 2016), https://www.boundless.com/economics/textbooks/boundless-economics-textbook/monopoly-11/monopoly-production-and-pricing-decisions-and-profit-outcome-71/marginal-revenue-and-marginal-cost-relationship-for-monopoly-production-268-12365/.

physical inputs, as well as some return to the labor of the producer needed to manufacture it.[107] The overall return to capital will be some rate, *r*.[108]

Consider now the consequence of relaxing the assumption of identical, zero risk tolerance. A circumstance necessary to support this assumption is the existence of surplus from the existing, noninnovative sector. Individuals (whether laborers or capital owners) with a positive risk tolerance are willing to risk a lower return, or perhaps even negative returns, in immediately succeeding periods in exchange for the prospect of a higher return in later periods, should their efforts bear fruit.[109] Individuals also must be willing to sacrifice leisure or current spending in favor of the prospect of higher returns from work or investment, respectively.[110] Some individuals in these groups merely seek higher returns from risk premiums but do not pursue innovation, while others seek higher returns by establishing an innovation sector.[111] (For most of the subsequent discussion, I disregard the first of these groups).

1. Nature of Returns in the Innovation Sector

All of the individuals entering the innovation sector, by definition, seek "new markets" (a term that includes better ways to satisfy existing markets)[112] and, consequently, leave the established labor and capital markets to some extent, reduce leisure or consumption, or both. Where they continue to work or invest in that market at less than "full-time" or with less than all capital, their departure is partial. As regards the reallocation of factors from the noninnovation to the innovation sector, the effects of this departure are twofold. On one hand, departing factors may or may not successfully innovate. If they succeed, they will experience, for a time, inframarginal returns because of the short-run excess of demand over supply of whatever successful innovation is supplied; these count as monopoly profits at the outset.[113] The returns will exceed *r*. If the effort is not successful, its returns will fall below *r* and may even be negative unless and until some other productive activity is found.[114] Nev-

107. *Id.*

108. *Return on Capital,* INVESTING ANSWERS FINANCIAL DICTIONARY, http://www.investinganswers.com/financial-dictionary/ratio-analysis/return-capital-3054 (last visited April 6, 2017).

109. *Risk-Return Tradeoff*, INVESTOPEDIA, http://www.investopedia.com/terms/r/riskreturntradeoff.asp (last visited April 6, 2017).

110. Stark, *supra* note 104, at 52.

111. For example, they may invest in activities the returns from which fluctuate on the basis of largely random, uncontrollable circumstances, such as rainfall in a given period.

112. Nizar Becheikh et al., *Lessons from Innovation Empirical Studies in the Manufacturing Sector: A Systematic Review of the Literature from 1993–2003*, 26 TECHNOVATION 644, 644 (2006).

113. *See* David A. Weisbach, *Ironing Out the Flat Tax*, 52 STAN. L. REV. 599, 605 (2000) ("The inframarginal return is the return on the investment above the normal rate of return (i.e. the rate of return on any additional cash invested). For example, monopoly profits are inframarginal returns.").

114. Leslie Feinzaig, *Investing in Disruptive Innovation*, FORBES (Aug. 13, 2008, 6:00 AM), http://www.forbes.com/2008/08/12/google-digitalequipment-disruption_leadership_clayton_in_lz_0813clayton christensen_inl.html. It is assumed for purposes of this discussion that regulatory mechanisms are

ertheless, on balance, returns from the innovation will exceed *r* since the prospect of increased returns is what motivates innovation.[115]

It is important to note the specific nature of returns to the innovation sector. As a general matter, returns to risk are considered to be lognormally rather than normally distributed. For example, the Black-Scholes model of option pricing assumes lognormal probability distributions for stock prices over time.[116] Lognormal distributions arise when the dependent variable (price, in this case) is a function of the mathematical product of a large number of identically distributed independent variables.[117] By contrast, a normal distribution arises when the dependent variable is a function of the mathematical sum of a large number of identically distributed independent variables.[118] Using a lognormal distribution for stock prices is appropriate because it corresponds to a normal distribution for expected *returns* on the stock, assuming continuous compounding of returns. That is, if continuously compounded returns on the stock price follow a normal distribution (so that actual rates of return are symmetrically distributed about the expected rate of return), then the stock price itself follows a lognormal distribution.[119] A lognormal distribution has a lower bound of zero, with a long "right-hand" tail reflecting the fact that there is, in principle, no upper bound to the stock price.[120] Figure 1 illustrates some sample lognormal distributions. The independent variable, *x*, is the input, which in this case is the return on an amount invested, assumed to be 1. The dependent variable is the probability of the return.

FIGURE 1: LOGNORMAL RETURNS–PROBABILITY DENSITY FUNCTIONS[121]

in place to ensure that innovators capture economic returns to innovation. These mechanisms would include the laws of patent, copyright, and trademark, or similar regimes.

115. Michael D. Makowsky & David M. Levy, *Increasing Returns to Scale, Price Dispersion, and the Distribution of Returns to Innovation*, 19 MACROECONOMIC DYNAMICS 334, 338 (2014).

116. JOHN C. HULL, OPTIONS, FUTURES, AND OTHER DERIVATIVES 300 (8th ed. 2011).

117. *See, e.g.*, Brian E. Smith & Francis J. Merceret, *The Lognormal Distribution*, 31 COL. MATHEMATICS J. 259, 260 (2000) (stating that if a distribution results from the product of a large number of independent variables, the log of the distribution will be normally distributed).

118. *Id.*

119. HULL, *supra* note 116, at 323–24.

120. *Id.* at 301.

121. Krihnaveddala, *Probability Distribution Functions of Log-Normal Distributions*, https://commons.wikimedia.org/wiki/File:PDF-log_normal_distributions.svg (last visited April 6, 2017).

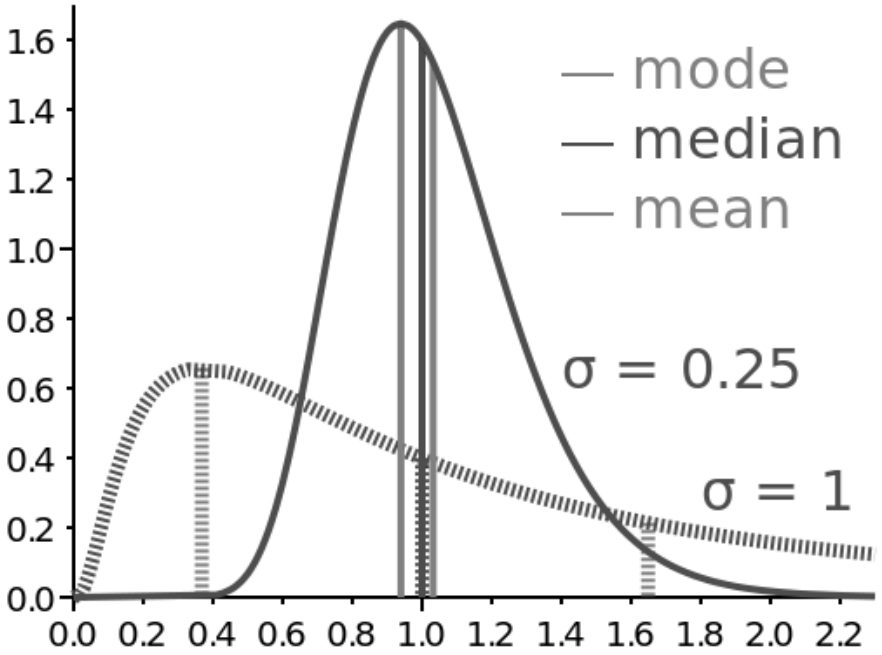
x

The parameters μ and σ respectively stand for the mean and standard deviation of the natural logarithm of the underlying variable (in the present context, return to underlying investment). By definition, the mean is the expected outcome.[122] The standard deviation expresses the extent of expected departures from the mean, so larger standard deviations imply that outcomes are more volatile.[123] Because these are standard deviations of logarithms of the underlying variable, however, the absolute value of a standard deviation on either side of the mean differs: it is lower for a negative standard deviation because the total range of possible outcomes below the mean is bounded at zero.[124] For outcomes above the mean there is no upper bound.[125] Correspondingly, the standard deviation of the underlying variable, mean return to investment, would be symmetrically distributed about that mean return.[126] If, for example, the mean return to investment were 3%, equal to 1.03, then the standard deviation of returns would be symmetrically distributed around 1.03 on an investment of 1.

The graphs in Figure 1 depict expected price changes of zero, corresponding to an expected rate of return of 1[127]—that is, if one invests $\$X$, one expects $\$X$ on average at the end of the period (mean equal to 1 on the continuously compounded return to investment). In practice, however, there is a tradeoff between nonvolatility and investment return, as depicted in Figure 2, for two possible investments. Because risk-bearing is costly, expected returns (μ) that are more volatile (corresponding to higher values of σ) will generally be higher than returns that are less so, with the risk-free rate being the baseline for the most nonvolatile investment possible.[128] Therefore, an increase in σ is associated with an increase in μ: assets or projects having a higher risk have a higher expected return.[129] Table 1 illustrates the principle by means of probability distributions for representative values of μ and σ.[130] The left-hand column indicates the return on an initial investment of 1, while columns (2) through (5) indicate the probability of receiving at least that return for a range of increasing parameters of μ and σ. As the values for the parameters increase, the likelihood of receiving higher returns increases as well.

122. *Id.* at 287–88.
123. *Id.* at 290.
124. *See id.* at 300–03.
125. *See id.*
126. *See id.*
127. Stated otherwise, if $ln(x) = 0$, then $x = 1$.
128. One mechanism for increasing returns to volatility is that risk aversion causes risk-averse investors to sell riskier assets and buy less risky ones. The migration to less-volatile investments drives down the yield to the less-risky assets and drives it up for the riskier ones. HULL, *supra* note 116, at 311–12.
129. *Id.* at 122.
130. No claim is made that the varying values of μ and σ in Table 1 correspond to the actual risk-return tradeoff, which is a function of risk aversion among investors.

FIGURE 2: LOGNORMAL RETURNS – PROBABILITY DENSITY FUNCTIONS, HIGH AND LOW σ[131]



131. Cmglee, *Comparison of Mean, Median, and Mode of Two Log-Normal Distributions*, WIKIMEDIA COMMONS, https://commons.wikimedia.org/wiki/File:Comparison_mean_median_mode.svg (last visited April 6, 2017).

TABLE 1: PROBABILITY DISTRIBUTIONS, SELECTED LOGNORMAL FUNCTIONS

(1) X	(2) Probability of x or less; μ = 0.09, σ = 0.3	(3) Probability of x or less; μ = 0.3, σ = 0.5	(4) Probability of x or less for μ = 0.69, σ = 1	(5) Probability of x or less for μ = 1.1, σ = 2
0.100	0.000	0.000	0.067	0.044
0.250	0.000	0.000	0.150	0.107
0.500	0.005	0.024	0.245	0.185
0.750	0.104	0.120	0.312	0.244
1.000	0.382	0.274	0.365	0.291
1.100	0.507	0.341	0.383	0.308
1.350	0.758	0.500	0.423	0.345
1.400	0.794	0.529	0.430	0.351
2.000	0.978	0.784	0.501	0.419
3.000	1.000	0.945	0.581	0.500
5.000	1.000	0.996	0.677	0.601
10.000	1.000	1.000	0.790	0.726
15.000	1.000	1.000	0.844	0.789
20.000	1.000	1.000	0.876	0.828
25.000	1.000	1.000	0.897	0.855
30.000	1.000	1.000	0.912	0.875
50.000	1.000	1.000	0.946	0.920
100.000	1.000	1.000	0.975	0.960

In general, the expected return rises as the mean and standard deviation rise.[132] The main point that is relevant for present purposes is that greater risk tolerance, corresponding to greater volatility, is associated with a greater expected return.[133] The significance of this point will become apparent in connection with the analysis of loss offsets, discussed below.

132. *See* HULL, *supra* note 116, at 122.

133. *Id.*; *see also Expected Return, Variance and Standard Deviation Of A Portfolio*, INVESTOPEDIA, http://www.investopedia.com/walkthrough/corporate-finance/4/return-risk/expected-return.aspx (last visited April 6, 2017).

2. *Effect on Noninnovation Sector*

Regardless of the success or failure of factors of production that move to the innovation sector, the fact of the departure will affect the return to factors that remain in the noninnovation sector. In the short run, the return to factors of production will increase because of the reduction in the supply of factors as the prospect of inframarginal returns in new markets siphons off resources to them.[134] Yet again, however, as the return to factors in the established market rises, so does the cost of goods there. The increase should offset, but not fully cancel, the effect of the rise in returns in the established market since one can expect, again in the short run, some reduction in overall demand as factors in emerging markets reduce consumption to increase their likelihood of success in the emerging market.[135]

From a welfare perspective, the addition of an innovation sector must be considered a positive development, assuming that risk-bearing factors meet with at least some success. This is so even though the sector's characteristics include the development of monopolistic firms in the short run.[136] By definition, the creation (or discovery) of a new, successful market means greater total social welfare, at least if we assume that consumers are generally rational in their consumption choices.[137] We can turn this point around by observing that if economic actors are generally rational, then a historical pattern of risk-bearing for the purpose of finding and developing new markets is welfare-enhancing because the pattern would arise only if there were a positive return to that risk-bearing sufficient to make it worthwhile for risk-takers to leave the established market with its certain return.[138] In short, the premium charged for risk-bearing, which takes the form of increased relative prices on established goods and explicit monopoly prices for some period of time on new goods, is worth it.

B. *Introduction of Taxes*

Having sketched out a simple model of innovation, the question becomes what happens when taxes are introduced into the model. In this Subsection I examine that question, beginning with ideal taxation and

134. *See* Magnus Henrekson & Tino Sanandaji, *Entrepreneurship and the Theory of Taxation*, 37 SMALL BUS. ECON. 167, 175 (2011) (discussing how supply of entrepreneurial effort, or innovative effort, is responsive or elastic to the availability of inframarginal returns)

135. *Id.* at 182 (assuming that there is an opportunity cost for postponed consumption that is associated specifically with entrepreneurial effort).

136. JOSEPH SCHUMPETER, CAPITALISM, SOCIALISM, AND DEMOCRACY 88–89 (1942). Schumpeter argued that the connection between successful innovation and monopoly pricing meant that perfect competition was not necessarily optimal from a welfare perspective. *Id.* The desirability of elevated profits in the innovation sector is addressed in more detail in Section II.B, above.

137. Graetz & Doud, *supra* note 1, at 348–50.

138. *See, e.g.*, Terrence R. Chorvat, *Ambiguity and Income Taxation*, 23 CARDOZO L. REV. 617, 620–22 (2002) (discussing expected-utility theory).

progressing first to simple second-best, which would be pure income taxation, and then to something approximating the actual income tax, which is a system that provides only limited relief for projects that sustain losses over their lifetimes. It becomes apparent that the limitation of loss offsets can have dramatic adverse efficiency effects on innovation.

1. Ideal Taxation

Under the standard model of ideal taxation, taxes either pay for pure public goods or finance redistribution, and in the following discussion I assume for ease of exposition that tax revenues are used solely for these purposes.[139] (Where tax revenues pay to supply other goods, the analysis changes somewhat, but not in ways that affect the basic argument[140]). Public goods are generally defined as goods that are nonrival and nonexcludable in consumption.[141] As previously explained, nonrival goods are those the provision of which does not diminish the quantity available to others; nonexcludable goods, on the other hand, are those the provision of which to any person implies their provision to everyone.[142] The feature of nonexcludability results in market failure because the producer effectively is unable to recoup the costs it incurs in providing the good.[143] The feature of nonrivalry means that there are no marginal costs to the consumption of the marginal unit of the good once it is supplied to the first consumer (or some fixed number of consumers).[144] As a consequence, once the good is supplied at a given level, an arbitrarily large quantity may be consumed at that level without requiring any further outlay.[145]

When taxes pay only for pure public goods or to finance redistribution, the ideal tax is generally thought to be lump-sum in nature.[146] A lump-sum tax is one that does not vary based on what the taxpayer does.[147] Lump-sum taxes are considered optimal in this setting because they do not cause the marginal cost of supplying the good to differ from the marginal cost of consuming it, and therefore they do not induce tax-

139. *See, e.g.*, J.A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38 REV. ECON. STUD. 175, 175 (1971).

140. In many cases, taxes pay for other types of goods or operate to internalize externalities ("Pigouvian taxes"). In theory, market instruments, together, possibly, with other legal rules, ought to be superior to broadly applicable taxes for nonpublic goods; in the case of externalities, taxes function as transfer mechanisms to align costs and benefits. Inclusion of an analysis of these types of taxes in the current discussion would add substantial complexity without changing the basic analysis, developed below.

141. Paul A. Samuelson, *The Pure Theory of Public Expenditure*, 36 REV. ECON. & STAT. 387, 387 (1954).

142. *See supra* Part I.

143. Samuelson, *supra* note 141, at 387.

144. *See supra* Part I.

145. William J. Baumol, *On Taxation and the Control of Externalities*, 62 AM. ECON. REV. 307, 312 (1972).

146. ROSEN & GAYER, *supra* note 47, at 333.

147. *Id.*

payers to allocate too few of their resources to the taxed good and too many to other goods.[148] A head tax, for example, must be paid regardless of the taxpayer's activity, which means the tax does not affect the relative prices taxpayers face in deciding how to allocate their resources; the marginal price of each good remains its marginal cost.[149] Thus, because the prices of goods continue to equal their marginal costs, no inefficiency arises from the tax.

Nevertheless, it is important to recognize that lump-sum taxes do result in changes to taxpayer behavior. Because taxpayers subject to any tax have fewer resources with which to satisfy their preferences, they must consume less, supply more work, or otherwise adjust their behavior in light of their lower budget. These adjustments are termed "income effects" to reflect the idea that they are the result of taxpayers' reduced incomes under the tax.[150] In short, because any tax reduces taxpayer wealth, and because consumption of a different bundle of goods is appropriate under a lower budget constraint, all taxes induce behavioral adjustments.[151]

Returning to the model of innovation just described, one can imagine at least three possible income effects from the imposition of a lump-sum tax burden. First, individuals may reduce present consumption or supply more labor generally. Except in the case of highly idiosyncratic risk-aversion functions, the supply of additional labor would be to the noninnovation sector.[152] Second, with respect to already-deployed factors of production, innovators may redirect resources from the innovation sector back to the noninnovation sector to ensure adequate resources to pay the tax. These redirections generally flow from risk-aversion.[153] Finally, innovators may reduce the riskiness of their investment in the innovation sector without moving resources out of the sector. For example, the innovation sector may consist of a variety of investment opportunities having different risk-to-return profiles. The presence of a lump-sum tax liability may cause an innovator to shift from a high-risk, high-return investment to a lower one. The shift is simply a more moderate version of shifting factors back to the noninnovation sector.

Again, it is important to bear in mind that these tax-induced effects, because they are income effects, are not inefficient; instead, they merely reflect the fact that any positive tax reduces private wealth, and changes of behavior due solely to reduced wealth are expected and rational.[154] As long as it is better for the government to have the tax revenue than for

148. *Id.* at 332.
149. *Id.* at 305.
150. *Id.* at 335.
151. *Id.* at 415–36.
152. Such a risk function could arise for an individual who needs to realize $*X* of income and for whom there is little or no utility difference between realizing slightly less than $*X* and much less than $*X*.
153. ROSEN & GAYER, *supra* note 47, at 185.
154. Weisbach, *Line-Drawing*, *supra* note 92, at 1652.

the money to remain in private hands, nothing is lost from income effects. Taxpayers experience reduced welfare from the tax, but, by assumption, the government generates greater welfare from the tax revenue than taxpayers lose by turning it over to the government.

2. *Pure Income Taxation*

For a variety of reasons, lump-sum taxes generally are considered unavailable.[155] For example, they may be considered regressive, or taxpayers may view them as imposing a kind of wage slavery if the tax burden causes individuals to work to satisfy the tax burden.[156] These perceptions may persist even if the extent to which "forced labor" arises is minimal and even if the government nullifies the regressive effects of the tax through transfer payments.[157]

Regardless of the merits of these objections, the efficiency effects of real taxes, such as taxes on income, become easier to understand in light of the lump-sum ideal. Unlike a lump-sum tax, real taxes alter the prices of taxed goods relative to other goods because the price of the taxed good increases even though the marginal cost to supply it does not.[158] If prior to the adoption of a tax on a good its price reflects the actual marginal cost to supply it, then after the tax the price exceeds the marginal cost.[159] When marginal price exceeds marginal cost, too little of the good is demanded relative to the actual social cost of supplying the good, and too much of nontaxed goods are demanded for analogous reasons. These shifts by definition represent efficiency losses, since the pretax allocation of resources was optimal.

As previously discussed, the cost arising solely from these substitution effects (as opposed to administrative and compliance costs) is called the "excess burden" or "deadweight loss" ("DWL") of the tax, and it increases roughly as the square of the rate of tax.[160] One may consider it the social cost of obtaining the tax revenue, expressed in terms of the reduction in total societal wealth (private plus governmental) resulting from a reallocation of resources in response to the fact that marginal costs to

155. This is an animating assumption of, for example, the optimal tax literature. *See, e.g.*, Mirrlees, *supra* note 139, at 175 (noting that the most readily observable indicator of wage rate is earned income, which is sensitive to tax rates).

156. For a discussion of these issues, *see generally* Daniel N. Shaviro, *Commentary: Inequality, Wealth, and Endowment*, 53 TAX L. REV. 397 (2000).

157. *Id.* at 416–20 (noting that the objections have resonance even if they may be hard to defend analytically); *see also* Mark G. Kelman, P*ersonal Deductions Revisited: Why They Fit Poorly in an "Ideal" Income Tax and Why They Fit Worse in a Far from Ideal World*, 31 STAN. L. REV. 831, 842 (1979) (providing an example of the intuition that endowment taxation is illiberal).

158. Typically, the burden of a commodity tax is shared by consumers and producers, who take a reduced profit even though the price increases. Consequently the price increase is less than the amount of the tax. The ratio according to which they share the burden depends on the relative elasticities of substitution faced by each. *See* ROSEN & GAYER, *supra* note 47, at 322.

159. *Id.* at 360.

160. *Id.* at 329.

consume no longer equal marginal costs to supply.[161] Hence, a broad base at a low rate is generally preferred to a narrower base at a high rate.

It is helpful to consider the effects of the introduction of a pure, flat-rate income tax, *t*, into the framework described above. A pure income tax is thought to introduce two significant distortions into taxpayer behavior.[162] It creates an incentive to replace work with leisure because leisure goes untaxed, and it creates an incentive to consume in the present rather than to save because income from savings is subject to tax whereas no further tax is due on income consumed.[163]

a. Labor Supply

Looking first at the labor-leisure decision, the substitution effect in the noninnovation sector is to reduce labor supply, as the tax reduces the value to workers of labor they furnish. Under the standard assumption of the declining marginal utility of income, the last dollars earned are worth less in utility terms than are inframarginal dollars.[164] Workers, therefore, no longer supply labor for the last dollars they otherwise would earn, assuming the reduction in wealth from the tax drives the after-tax value of the last dollars earned below the value of the leisure they would enjoy if they did not work.

The effect in the innovation sector is somewhat different. For innovators who are somewhat resource-constrained, adoption of the tax reduces risk tolerance. Consider that an innovator can assume risk only if she has sufficient resources to cover the cost of some period of unsuccessful innovation. Therefore, for relatively less risk-tolerant individuals, the tax should cause a shift of labor back to the noninnovation sector. The tax also could have the countervailing effect, as workers expend more labor to realize higher returns. Both of these effects, however, are income effects. They do not result from a reduction in the value of labor *vis-à-vis* leisure but from the absolute reduction in wealth caused by the tax. Whether in any particular case a worker shifts labor supply to the noninnovation sector or supplies more labor to the innovation sector is a function of the worker's risk aversion and the availability of additional resources to devote to innovation. In either case, no inefficiency results from the reallocation.

For relatively more risk-tolerant individuals, the tax should cause a simple reduction in labor supply for reasons analogous to those in the noninnovation sector. As in the case of the noninnovation sector, the re-

161. *Id.* at 338. DWL tends to increase as the square of the tax rate because an increase in rate causes an increase in both the number of substitutions made and the marginal cost of each substitution. Weisbach, *Line-Drawing, supra* note 92, at 1656.

162. ROSEN & GAYER, *supra* note 47, at 417.

163. *Id.*

164. Richard A. Easterlin, *Diminishing Marginal Utility of Income? Caveat Emptor*, 70 SOC. INDICATORS RES. 243, 243–44 (2005) (noting caveats to the general principle but acknowledging its widespread acceptance in the literature).

duction in labor supply in response to the tax is an inefficient substitution effect. For each individual, how much a type of effect appears—income in the form of a shift or intensification of labor supply, or substitution in the form of a reduction in labor supply—depends on the shapes of the relevant indifference curves.[165]

b. Capital Investment

In general, the imposition of a tax, *t*, on fixed, positive returns to capital investment can cause a reduction in savings because the tax reduces the value of the future income stream from savings. This would be a substitution effect, as it results from the change in after-tax return to capital investment, driving some portion of the value (in utility terms) to the taxpayer of future consumption below the value of additional current consumption.[166] The tax also can have an income effect. Where the tax causes returns to drop below an investor's target, the response is to increase investment (typically financed by a reduction in consumption) to make up for the lost income.[167]

Things may be somewhat different where capital returns fluctuate. Under a pure income tax, taxpayers receive full loss offsets, which means taxpayers who experience a net loss during the tax period receive a subsidy (a negative tax) from the government that is analogous to the tax they pay on positive income.[168] If, for example, *t* is a flat 30%, then taxpayers who lose 10% during the tax period receive a payment from the government of 3%. The payment reflects the fact that the taxpayer previously paid tax on the capital for which the taxpayer does not derive a benefit. An extensive literature has examined the effects of full loss offsets under a pure income tax and, as a general matter, has concluded that the tax is at least somewhat unable to burden the returns to risk-taking.[169] The reason is that, under a full-loss-offset regime, the investor can respond to the tax on the risky return by increasing his investment in assets delivering risky returns. When the asset makes money, the enlarged return covers the associated tax liability, and when it loses money, the enlarged deduction likewise returns the taxpayer to the loss position he would have experienced in the nontax world. For example, if the investment will either pay 3 or lose 2 on an investment of 10, the investor in a nontax world will end up with either 13 or 8. If a 30% tax is adopted, the win drops from 3 to 2.1 and the loss drops from 2 to 1.4. The taxpayer

165. ROSEN & GAYER, supra note 47, at 417.
166. *Id.*
167. *See* Weisbach, *Line-Drawing, supra* note 92, at 1656 (discussing the "income effect").
168. *See* Surrey, *supra* note 87.
169. Domar and Musgrave initiated the modern literature. Evsey D. Domar & Richard A. Musgrave, *Proportional Income Taxation and Risk-Taking*, 58 Q. J. ECON. 388, 391 (1944). For a recent general overview, see Wolfgang Buchholz & Kai A. Konrad, *Taxes on Risky Returns – an Update*, 1 (Max Planck Inst. for Tax Law & Pub. Fin., Working Paper No. 2014-10, (July 2014), https://papers.ssrn.com/sol3/papers.cfm?abstractid=2469268.

can negate the effect of the tax by increasing the investment by $t/(1 - t)$ times the initial investment, or 43% in the case of a 30% tax. On an initial investment of 14.29, a win yields 18.57 pretax, which produces a pretax gain of 4.29 and an after-tax gain of 3. The same initial investment yields 11.43 on a loss, or a pretax loss of 2.87 and an after-tax loss of 2, just as before the tax was in effect.

If taxpayers can fully negate the tax on risky returns, there is literally no income or substitution effect from the tax; the tax on risky returns becomes a nullity.[170] There is, however, substantial reason to question whether taxpayers can "scale up" their investments in this way to avoid the tax in the context of opportunities for innovation.[171] In the first place, the scaling-up argument depends critically on taxpayers' capacity to exchange riskier for less-risky projects. Therefore, if capital may not be readily redeployed for this purpose, the increased appetite for risk may not be matched by opportunities to indulge it, and taxpayers cannot effectively shift the tax back to the government.[172] Secondly, and more importantly, the scaling-up argument, however much truth it carries with respect to returns to systemic risk, generally does not apply to "inframarginal returns," or returns from special investment opportunities that innovation supplies, which by definition are limited in scope.[173] As contrasted with opportunities that may be available to invest in garden-variety forms of risk-taking, the opportunities available to invest in a new technology or other innovation are inherently limited by the nature of the investment. Once an amount sufficient to produce the quantity of the good that satisfies consumer demand has been invested, there is no more reason to invest in the innovation.

In what follows, I assume that scaling opportunities, even under a pure or "normative" (as it is sometimes termed) income tax, are extremely limited for investments in innovation, most of which are, by nature, opportunities limited in scope and carrying with them inframarginal returns (albeit subject to substantial risk). Consequently, the income and substitution effects of a normative income tax on the innovation sector are as described above. In the main, the income effect is a shift to the noninnovation sector, and the substitution effect is a reduction in amounts invested in capital. Again, this analysis assumes full loss offsets are available.

170. Buchholz & Konrad, *supra* note 169, at 2.

171. *See generally* Theodore S. Sims, *Capital Income, Risky Investments, and Income and Cash Flow Taxation*, 67 TAX L. REV. 3, 24–25 (2013) (arguing that scaling up will be limited unless taxpayers exhibit constant risk aversion).

172. *Id.* at 17–18.

173. *See, e.g.*, David A. Weisbach, *The (Non)Taxation of Risk*, 58 TAX L. REV. 1, 19–21 (2004) (noting that it is impossible to scale up investment in assets returning inframarginal returns for the reason stated in the text).

3. *Real Income Taxation*

A serviceable model of the actual income tax is further removed from an ideal, lump-sum tax in that it requires loss limitations.[174] In assessing the effects of the actual tax system on innovation, I therefore begin with the income tax described above, except that in place of full loss offsets I assume imperfect loss offsets. In simple terms, under imperfect loss offsets, deductions for negative income are not allowed; capital losses are generally deductible only against capital gains, with a carryover regime for unused losses; and net losses for the tax year are suspended, likewise available for possible use in a prior or future year, but only under restricted conditions.[175] All of these features are found under the actual income tax.[176]

The ordinary- and capital-loss limitations work similarly, though their rationales are not identical. The capital-loss limitation generally permits capital losses during the taxable year to offset only capital income, except for up to $3,000 of ordinary income for noncorporate taxpayers.[177] Excess losses are carried to another year—typically forward, though corporations may carry back up to three years.[178] The more general loss limitation applies when otherwise allowable deductions exceed gross income. In that case, of course, no tax is due, but no negative payment from the government is forthcoming either.[179] Instead, the taxpayer has a net operating loss ("NOL"), approximately equal to the excess, that may be applied as a deduction to offset the same taxpayer's positive income in another tax year.[180]

The principal purpose of the capital-loss limitation appears to be to reduce taxpayers' ability to control the timing of income and loss for tax purposes; this kind of control enables taxpayers to reduce reportable taxable income relative to true economic income.[181] The mechanism for

174. All income-tax systems place some limitation on the availability of loss offsets. *See* PRICE WATERHOUSECOOPERS, *Worldwide Tax Summaries: Corporate Taxes* (2014), https://www.pwc.com/gx/en/tax/corporate-tax/worldwide-tax-summaries/assets/pwc-worldwide-tax-summaries-corporate-2014-15.pdf.

175. *See also* Alan J. Auerbach & James M. Poterba, *Why Have Corporate Revenues Declined?*, *in* 1 TAX POLICY AND THE ECONOMY 11–12 (Lawrence H. Summers ed., 1987) (discussing the relationship between imperfect loss offsets and average tax rates).

176. The nonrefundability of overall losses follows from the language of § 1 (individuals) and § 11 (corporations). I.R.C. §§ 1, 11 (2012). Carryovers of net operating losses are provided for in § 172. *Id.* § 172. The capital loss limitation and carryover provisions are in § 1211 and § 1212. *Id.* §§ 1211, 1212. In addition, § 382 sets out extensive limitations on the use of one taxpayer's loss by another taxpayer. *Id.* § 382.

177. *Id.* § 1211(a) (corporations), (b) (other taxpayers).

178. *Id.* § 1212(a) (corporations), (b) (other taxpayers).

179. *Id.* § 1.

180. *Id.* § 172(a). Not all deductions available for the purpose of computing taxable income are available for the purpose of computing an NOL. For example, no deduction is available for personal exemptions, and business and personal expenses are segregated to some extent. *See id.* § 172(d)(3)–(4).

181. Daniel Shaviro, *Selective Limitations on Tax Benefits*, 56 U. CHI. L. REV. 1189, 1191–92 (1989).

control is the realization rule, or the principle that capital income or loss is not taxed until the taxpayer disposes of the asset that has appreciated or declined in value.[182] In the context of capital income and loss, the problem is fairly straightforward. Taxpayers may own positions that have risen in value and positions that have declined in value. A taxpayer can realize the losses while holding on to the gains, thereby reducing taxable income even though the taxpayer's real economic wealth includes the untaxed appreciation in the gain positions. Without a capital-loss limitation, losses would effectively be taxable on an accrual basis, while gains remain taxable on a realization basis, negating the tax on the time value of money and thereby moving towards consumption taxation. The capital-loss limitation restricts taxpayers' ability to move toward such a mixed-base regime for capital income and loss.[183]

The reasons for the analogous prohibition on deductibility of losses when total income (not merely capital income) is less than zero are somewhat more obscure. The worry that taxpayers will manipulate their ordinary income in analogous fashion to their capital income seems overstated. Unlike capital income, taxpayers generally have relatively less control over the timing of realization of ordinary income or expenses.[184] Therefore, taxpayers reporting negative taxable income likely do not have associated unrealized economic income. Nevertheless, income-tax systems generally restrict the availability of deductions for negative income, typically by NOL-type provisions.[185] Among the additional possible reasons for denying full loss offsets may be worries about taxpayer fraud or the view that an income tax normatively applies only to positive income.[186] Another is that it is unclear what the rate of negative tax should be. If the positive tax rate were flat, a subsidy at the same rate seems, at least intuitively, plausible but given that positive rates are graduated, it becomes harder to justify any particular schedule for negative income.

As Nussim and Tabbach have noted, none of these rationales is fully persuasive.[187] Fraud is possible regardless of income level, though perhaps there is something to the thought that one who has demonstrable positive income from a trade or business is not as likely to defraud the government as one who has no income at all—the latter possibly having no trade or business in the first place.[188] Similarly, the fact that the shape

182. I.R.C. § 1001(a).

183. *See, e.g.*, Robert H. Scarborough, *Risk, Diversification and the Design of Loss Limitations Under a Realization-Based Income Tax*, 48 TAX L. REV. 677, 685–92 (1993).

184. Chye-Ching Huang, *Recent Studies Find Raising Taxes on High-Income Households Would Not Harm the Economy*, CENTER ON BUDGET AND POLICY PRIORITIES (Apr. 24, 2012), http://www.cbpp.org/research/recent-studies-find-raising-taxes-on-high-income-households-would-not-harm-the-economy.

185. Nussim & Tabbach, *supra* note 21, at 1525.

186. *Id.* at 1561–65.

187. *Id.*

188. *Id.*

of the negative tax schedule is not immediately clear does not explain the policy choice to deny negative taxes entirely, just as the facts that varying rates from year to year or of graduated rates do not defeat the application of NOL deductions in other taxable years at whatever rate happens to apply.[189] Finally, the denial of full loss offsets seems to require the questionable notion that there is something special about zero as a line below which expenses do not count as a reduction to income.[190]

At all events, the implications of the limitations on loss offsets are significant for innovative activity. Most innovation is conducted by taxpayers through the regular (not "S") corporate form, which means that gains and losses at the entity level generally remain separate from the gains and losses of the entity's owners.[191] Consequently, when the corporation experiences a net loss, the loss is not available to shareholders to offset their income from other sources. Therefore, they (through the corporation) are remitted to the NOL regime to obtain relief for net losses.

The NOL regime is relatively ill-suited for mitigating economic losses. Consider the simple case in which the firm conducts one trade or business that is entirely engaged in innovative activity. The firm faces a theoretical schedule of pretax returns from its activities that reflects the possibilities of having positive or negative outcomes with varying probabilities. For example, the firm may judge its activity to be moderately, but not excessively, risky, such that it will "win" half of the time and lose the other half of the time, with the expected net gain greater than the riskless rate to make it worthwhile to the firm's owners in light of their risk preferences. Or, the firm may have the option of investing in a riskier project having a lower probability of a win, but enjoying a larger gain if it does win. Under a flat-rate tax with no loss offsets, the expected return drops by more than the rate of tax times the pretax return.

4. *Incentive Effects*

Imperfect loss limitations lead to a number of immediate substitution effects; these effects, in turn, have secondary effects. Immediate effects may include driving capital out of the innovation sector to avoid the generation of unusable losses; driving investment diversification in the innovation sector to make otherwise unusable losses usable; and finally, and in tension with the first two effects, driving investors in the sector to make riskier investments (rather than to reduce risk by exit or diversification) to take advantage of the lognormal distribution of returns from innovation.[192] Secondary effects may include reduced rates of return in

189. *Id.*

190. *Id.* at 1526–27, 1566 (discussing how losses are not refundable under any system).

191. *See* I.R.C. § 11 (2012) (applying tax to corporate income). Startup ventures generally take the "C corporation form," despite the apparent tax disadvantages of doing so. *See* Susan C. Morse, *Startup Ltd.: Tax Planning and Initial Incorporation Location*, 14 FLA. TAX REV. 319, 343 (2013).

192. *See supra* Subsections III.A.1–2.

the noninnovation sector, enhanced principal-agent losses in the innovation sector, and the creation of systemically dangerous risk.[193] The following discussion briefly describes these primary and secondary effects.

a. Shift out of the Innovation Sector

The prospect of high taxes in the innovation sector can drive capital back to the noninnovation sector, or it can induce greater current period consumption. The latter can result up to the extent that the creation of an innovation sector drew capital from current consumption in the pretax world. To the extent capital moves back to the noninnovation sector, returns to capital in that sector drop while returns in the innovation sector rise and become more volatile.[194] The latter effect arises whether capital moves to the noninnovation sector or is currently consumed.[195] To the extent labor and capital are complements, labor will also be shifted back to the noninnovation sector, suppressing wages there as well.[196]

There is substantial evidence of the impact of taxation on exit from the innovation sector. Gentry and Hubbard observe a significant correlation between increasing marginal tax rates and the reduction of entry into the entrepreneurial sector,[197] noting as well that imperfect loss offsets have substantially the same effect as a progressive rate structure.[198] More generally, Feldstein notes the impact of progressive taxation on labor supply, which may have correlates in the decision either to deploy capital in sectors with lower tax rates (*i.e.*, in the present setting, where overall losses are less likely) or to consume, rather than invest, capital.[199]

b. Increased Diversification

A second method by which to reduce the effective tax rate is to diversify. Firms can do this by running more than one innovative business under the same corporate umbrella, or by combining innovative activity with noninnovative activity under the umbrella. As contrasted with investment portfolio diversification, which has relatively little impact on the day-to-day operations of the firms in the portfolio, pooling active

193. *Id.*

194. The phenomenon of adjustments to returns as factors of production move is a standard assumption of economic theory. *See, e.g.*, Arnold C. Harberger, *The Incidence of the Corporation Income Tax*, 70 J. POL. ECON. 215 (1962) (arguing that, in a closed economy, all capital bears the incidence of an income tax nominally assessed only on returns to corporate capital, because the tax causes factors of production to migrate from the corporate to the noncorporate sector).

195. *See id.*

196. *See id.*

197. Gentry & Hubbard, *supra* note 26, at 286.

198. *Id.* at 283.

199. Martin Feldstein, *The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act*, 103 J. POL. ECON. 551, 553 (1995); *see also* Gentry & Hubbard, *supra* note 26, at 286 (noting that the implicitly progressive tax schedule on capital due to incomplete loss offsets may have the same dampening effect on entry into entrepreneurial activity that explicitly progressive rates have on labor income).

business activities at the firm level is likely to impose social costs.[200] It causes the allocation of effort to individual enterprises to diverge from a nontax-affected optimum; it also may lead to adverse scaling effects. For example, two firms may decide to combine in order to ensure the usability of future losses realized by either firm. As a general matter, past losses of either firm will be unavailable to offset the income of the other firm.[201] To effectuate the combination, there may be changes in firm governance, reallocation of duties among entrepreneurs, or redirection of firm resources among firm projects. These adjustments, taken on their own, are inefficient if they are tax-motivated. The large body of scholarship on the theory of the firm seeks to identify the reasons why firms exist when, in theory, market transactions between individuals could provide the same benefits.[202] The basic theory is that transaction costs make market contracting too expensive relative to the coordinated arrangement among firm members up to a certain level of transacting.[203] Nevertheless, the fact that all business is not conducted in one firm indicates that, at some point, countervailing considerations make markets more efficient than combination, even in light of transaction costs. A tax rule that drives otherwise inefficient combinations therefore imposes social costs by putting within the firm operations that should be conducted by separate firms.

The proposition that the more fulsome tax reduction regime that accompanies diversification would encourage business combinations is fairly intuitive.[204] At least one paper corroborates the hypothesis that taxes can drive business combinations, noting in a study of more than 42,000 firms across more than forty countries that higher statutory corporate tax rates are correlated with larger firm size.[205] Indeed, the proposition nearly follows purely as a logical matter from Congress's perceived need to enact § 382 as a method to reduce trafficking in losses (as well as § 384, which applies similarly to limit a corporation's use of gains from an acquired business to soak up the acquiring corporation's preacquisition losses).[206] Section 382 prevents an acquiring company from using most of the already-realized losses of a target company to offset its own taxable

200. Feldstein, *supra* note 199, at 554–55.

201. I.R.C. § 382 (2012) (limitation on acquiring corporation's use of target's NOLs); *id.* § 384 (limitation on an acquiring corporation's use of its preacquisition losses to offset target's built-in-gains).

202. The seminal paper is Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937).

203. *Id. at* 390–91.

204. *See, e.g.*, Roberta Romano, *A Guide to Takeovers: Theory, Evidence, and Regulation*, 9 YALE J. REG. 119, 134 n.60 (1992).

205. *See, e.g.*, Peter Egger et al., *Heterogeneous Tax Sensitivity of Firm-Level Investments* 40 Tables 2, 3 (Inst. for Advanced Studies, Vienna and Univ. of St. Gallen, Working Paper No. 306, 2014).

206. "If . . . carryovers can be transferred in a way that permits a loss to offset unrelated income, no legitimate averaging function is performed. With completely free transferability of tax losses, the carryover provisions become a mechanism for partial recoupment of losses through the tax system." JOINT COMM. ON TAXATION, GENERAL EXPLANATION OF THE TAX REFORM ACT OF 1986, H.R. Doc. No. 3838, at 294 (1987); *see also* I.R.C. § 384 (2012). The text discusses § 382, but analogous considerations apply to § 384.

income, but only when the losses are realized preacquisition.[207] Because the provision simply limits firms' capacities to take unrelated losses *ex post*, however, its effect is to encourage moving the acquisition forward in time, before the target's loss is realized, so that the unrelated loss is available to the acquiring corporation once the loss materializes.

c. Shift to Higher Expected-Payoff Activities

As previously discussed, the taking on of riskier projects is in tension with the strategies of exit and combination as ways to reduce tax liability strategy. This strategy can be effective because of the lognormal distribution of returns from risk-taking. Recall that a lognormal distribution retains a median return at or near unity for most parameters, but that returns distributed about the mean have a lower bound of 0 and no upper bound.[208] Therefore, loss outcomes (where losses are understood as measured relative to the expected return), though as frequent as gain outcomes, are grouped in a narrower region so that the mean outcome will exceed the median—gains generally exceed losses. The extent of the asymmetry, however, depends on the parameters of the distribution. Distributions that are more skewed to losses (in terms of the modal, or most likely, outcome) feature a relatively higher expected loss for all loss outcomes than do less-skewed distributions,[209] but a much higher expected gain for all gain outcomes, again, as compared to less-skewed distributions.[210] The result is that the ratio of the expected gain for gain outcomes to expected loss for loss outcomes is larger than for distributions having less-skewed parameters; in other words, the mean return rises. Perhaps counterintuitively, in a world of no loss offsets, this ratio can become much larger, pushing investors to riskier projects.

As an illustration, let us suppose that a company faces two potential projects. On an investment of 100, the less-risky project has possible payouts of 170 on a win and 50 on a loss, before any tax. That is, the investment either wins 70 or loses 50. Assume the payouts have equal probabilities so that the expected payout is 110.[211] On the same investment, the riskier project has expected payouts of 220 and 30 (corresponding to wins of 120 and losses of 70), again with equal probabilities, for an

207. I.R.C. § 382(a).

208. *See supra* Figures 1 and 2.

209. Compare the mean loss for loss outcomes of the less-volatile (lower sigma value) investment with the mean loss for loss outcomes of the more-volatile (higher sigma value) investment in Figure 2.

210. Compare the mean gain for gain outcomes of the less-volatile (lower sigma value) investment with the mean gain for gain outcomes of the more-volatile (higher sigma value) investment in Figure 2. Finally, note that the increase in expected return for mean gains of the more-volatile investment outweighs the increase in expected loss for mean losses of that investment. In other words, a larger expected increase on gain outcomes than reduction on loss outcomes arises with increasing volatility.

211. The expected payout is given by the gain or loss times the probability of each, or $0.5*70 + 0.5*(-50) = 10$.

expected payout of 125 before the tax.[212] Under a 30% income tax with full loss offsets, the expected payouts drop to 107 and 117.5, respectively. The ratio of the expected returns on the two investments remains the same before and after the tax: two-and-a-half to one. The income effect of the tax may be to induce shifting of resources from the high-risk to the low-risk project; under the assumption of declining marginal returns to wealth, the shift will not be to a higher-risk project.

Now consider the effects of the adoption of the same tax, except that the tax has no loss offsets. The expected payout on the moderately risky investment drops to 99.5;[213] that is, the expected return becomes negative. The expected payout on the riskier project drops to 107.[214] The reason for the disproportionate drop in expected returns is that the greater symmetry between gains and losses for the less-risky project means that the denial of loss offsets raises the tax on a larger percentage of the expected return than it does on the riskier project.

As described above,[215] the literature has generally assumed the consequence of denying loss offsets would be either a reduction in risk-taking or combination with other investments,[216] but the considerations developed here indicate that a third response may be *greater* risk-taking. Which course of action is more attractive to the investor in our example depends on the investor's risk profile and on the expected returns under available alternative strategies, but it is not clear that less aggressive risk-taking would always be chosen. For example, combining with another entity may create efficiency losses (through, for example, principal-agent problems) so that even though the prospect of overall losses is eliminated, the expected return becomes unacceptably low (say 5 rather than 7 under full loss offsets). Similarly, the return in the riskless sector may be enough lower that redeploying capital there remains unattractive, even though no implicit tax from losses is possible (because the riskless rate is nonnegative). In short, it seems likely that in certain cases, the risk-return profiles of available alternatives would result in additional risk-taking.

Critical to an assessment of the extent to which loss limitations in practice lead to riskier investments is some idea of the actual relationships between risk and return. We would like to know, for example, how a doubling of volatility from 0.2 to 0.4 affects the expected return on the underlying investment relative to a given riskless rate. The answer to the question, by its nature, is somewhat elusive because, among other things,

212. The expected payout is 100 + (0.5*120 + 0.5*(-70)).
213. The expected payout is 100 + (0.5*70*0.7 + 0.5*(-50)).
214. The expected payout is 100 + (0.5*120*0.7 + 0.5*(-70)).
215. *See supra* Part II.
216. *See, e.g.*, Gentry & Hubbard, *supra* note 26, at 283 (noting that imperfect loss offsets can discourage entry); Robert H. Scarborough, *Risk, Diversification and the Design of Loss Limitations Under a Realization-Based Income Tax*, 48 TAX L. REV. 677, 679 (1993) (noting the availability of diversification as a tool to manage risk and reduce the disadvantage of loss limitations).

underlying volatility itself is only measurable *ex post*, which means that volatility measures include random fluctuations in prices that may not reflect actual volatility.[217] Nevertheless, some data are available, as set forth in Table 2.

TABLE 2: ANNUALIZED CPI RELATIVE RISK-RETURN ESTIMATES: 1973-2015[218]

Asset Classes	**Return (%)**	**Std. Deviation (%)**
Fixed Income	0.8	1.2
Intermediate Term Bond	3.1	4.5
Commodities	3.3	19.7
International Gov't Bonds	4.5	10.2
Long Term Bond	4.7	10.0
High Yield Bonds	5.8	8.3
International Equity	6.4	17.4
Large Cap Equity	6.9	15.5
Mid Cap Equity	9.0	17.5
REITs	9.5	18.9
Small Cap Equity	10.7	21.3
Emerging Market Equity	11.2	23.6

Table 2 is based on data compiled from 1973 to 2015. The relationship between risk and return set out there illuminates the tax advantage of taking on riskier projects.[219] In general, higher returns are closely correlated with greater volatility. A project with a relatively modest risk profile is likely to produce an expected gain for favorable outcomes that is only slightly larger than the expected loss for unfavorable ones on a pretax basis, whereas a riskier profile will produce a larger excess. From an *ex ante* perspective, for a relatively low risk-free rate, the latter project has a lower expected tax rate. If innovators respond to the tax disadvantage by taking on riskier projects as a way to reduce the expected tax rate, the overall effect is to skew investments in favor of riskier projects.

Although it is hard to quantify the extent of the effect, the proposition that excess risk-taking can be costly is well established. For example,

217. *See* Turan G. Bali & Armen Hovakimian, *Volatility Spreads and Expected Stock Returns*, 55 MGMT. SCI. 1797, 1797 (2009).

218. RICHARD MICHAUD ET AL., FI360 ASSET ALLOCATION OPTIMIZER: RISK-RETURN ESTIMATES 5 Table 2 (2016) (document on file with author) (reproduced by permission of authors).

219. The correlation between return and risk is quite high, with *r-squared* of 0.814.

research on the tax incentive to finance new projects with debt rather than equity (because of the deductibility of interest payments but not of dividends) has largely concluded that excess leverage has been a significant driver of recent cycles of boom and bust, with estimates in the early 1970s of annual deadweight loss equal to 10% of corporate tax revenues.[220] Similarly, the 2008 financial crisis has been widely blamed on excess risk-taking.[221]

Using the data in Table 2, it is possible to identify actual situations in which it could be rational to take on more risk in response to the absence of loss offsets. Begin by noting that Table 2 identifies a reasonably close correlation between arithmetic return (investment return rather than asset price) and volatility. The best-fit line for the data is:

(1) $f(x) = 1.758x + 2.890$

where x is return and $f(x)$ is volatility. Equation (1) has an r-squared value of 0.662. Speaking very roughly, a 1% increase in returns is associated with a 1.76% increase in volatility, after accounting for a fixed rate at zero risk of 2.89. There is, however, one substantial outlier in the data, which is the risk-reward relationship for commodities (entry three on Table 2). It is unclear why the relationship in this entry differs so substantially from that in the others, but if one assumes that it is anomalous and for that reason computes the best-fit line without it, the equation becomes:

(2) $f(x) = 2.098x - 0.355$

with an r-squared value of 0.922. This very high value of r-squared indicates a substantial correlation over the observed range between risk and return.

More specifically, for present purposes, one can say that the data suggest a roughly 2% increase in volatility for every 1% increase in returns—albeit with many caveats.[222] This relationship suggests that if limited loss offsets reduce returns enough, investors may migrate to projects having greater pretax volatility rather than less because the reduction in expected return will be sufficiently less.

Specifically, at the limit (where the riskless rate is low and there are no loss offsets), the *ex ante* tax rate approaches $2t$ on investments carrying modest risk because gains and losses are distributed nearly symmetri-

220. Benjamin M. Friedman, *Introduction*, *in* THE CHANGING ROLES OF DEBT AND EQUITY IN FINANCING U.S. CAPITAL FORMATION 1, 12 (Benjamin M. Friedman ed., 1982).

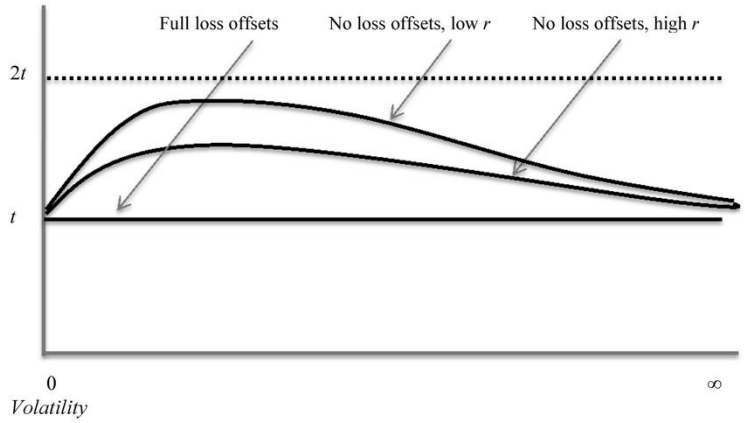
221. *See, e.g.*, Sewell Chan, *Financial Crisis Was Avoidable, Inquiry Finds*, N.Y. TIMES (Jan. 25, 2011), http://www.nytimes.com/2011/01/26/business/economy/26inquiry.html.

222. The data correlate volatility and return for publicly traded investments, not private investments, and operate over a relatively narrow range of possible risk-return relationships. In addition, they are taken over a single, approximately forty-year period that may not be broadly representative.

cally about the riskless rate—the distribution is nearly normal. Thus, the average gain, *G*, is taxed at *t*, and the equally likely average loss, -*G*, is taxed (subsidized) at 0. The firm therefore faces a choice: it can move to less-risky investments such that the *ex ante* tax rate moves closer to *t* (because the likelihood of negative returns drops), or it can move to more-risky investments such that, again, the *ex ante* tax rate moves closer to *t* (because the size of expected negative returns relative to that of expected positive returns diminishes). If the expected return on the riskier investment is more than 2% greater per 1% increase in volatility, then for the marginal investor it would be worthwhile to move to the riskier project rather than to the less-risky one. Note that this outcome becomes more likely the closer the risk-free rate is to zero, as it has been for most of the last decade.

Figure 3 depicts the relationship just described between *ex ante* tax rates and volatility for a given flat statutory rate, *t*, when loss offsets are denied. The rate rises from zero to a maximum at some point of moderate volatility, declining thereafter as volatility increases. The height of the "hump" depends on the riskless rate, which can be considered a parameter for this purpose. If the riskless rate is high, the hump will be lower because relatively nonvolatile investments will be unlikely to generate losses, even if they have a 50% chance of generating returns less than the riskless rate. If the riskless rate is low, the hump will be higher, with a maximum expected rate near 2*t* for investments having a relatively symmetrical distribution of gains and losses (that is, a near-normal distribution).[223] When the riskless rate is low, one can expect many investors to jump to riskier projects, creating substantial efficiency losses.

FIGURE 3: *EX ANTE* TAX RATES FOR STATUTORY TAX RATE T, RISKLESS RATE R *EX ANTE* TAX RATE



223. The analysis suggests that, in the current environment of very low riskless rates, which has persisted for more than a decade, excess risk-taking has been higher.

C. Pairing Full Loss Offsets with Higher Rates on Innovation

Observe that the expected tax schedule in Figure 3 has precisely the opposite shape from the optimal schedule, taking into account the declining marginal utility of money. Investments yielding a low return ought to be taxed less than investments yielding a high return, but in fact investors that are more risk-tolerant end up more lightly taxed on an *ex ante* basis, even under a nominally-flat rate. By contrast, under full loss offsets, the *ex ante* tax rate would equal the flat statutory tax rate for all volatility levels and the efficiency loss described above disappears.

Note, however, that it is possible to do even better (from an efficiency standpoint) than what is obtained under a full loss offset regime that produces a constant expected tax rate, t. For reasons explained below, innovation is unlikely to be sensitive to high tax rates on high income *ex post* outcomes. Consequently, pairing full loss offsets with higher expected rates on such outcomes can collect greater dollar revenue in the innovation sector than can a full loss offset regime with a flat tax, but at little or no additional efficiency cost. The increase in revenue in turn permits a reduction of tax rates in the noninnovation sector, thereby creating a second welfare benefit, as the deadweight loss of taxes in that sector is reduced.

The conclusion that high tax rates on successful risky projects are unlikely to affect pretax decisions is a standard result of optimal income tax theory.[224] Begin with the case of fixed returns in the noninnovation sector. Optimal tax theory derives a schedule of declining marginal tax rates as income increases above a relatively modest absolute level because of the declining marginal utility of the next dollar.[225] A taxpayer who has earned $\$n$ determines whether to continue to work based on the marginal utility derived from the $(n + 1)^{th}$ dollar. A tax on dollars reduces the utility derived from earning that dollar, and since for large values of n the utility from the next dollar is low, the tax must be low to avoid deadweight loss (namely, the tax-induced decision not to continue to work).[226] Similar principles hold in the case of investment income: the question is whether to invest or to spend currently (consume). The tax reduces the utility of the next dollar earned on saving but not on consuming, and as income rises the after-tax utility of returns to saving declines

224. Ravi Kanbur et al., *Moral Hazard, Income Taxation and Prospect Theory*, 110 SCANDINAVIAN J. ECON. 321, 325 (2008) (noting that the standard result of optimal tax theory that tax rates should be higher under greater risk aversion).

225. These points are developed in N.H. Stern, *Optimal Taxation*, *in* THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS (John Eatwell et al. eds., 1987), http://www.dictionaryofeconomics.com/article?id=pde1987_X001618&goto=O&result_number=2650#header.

226. *Id.*

as well. The tax on capital income, therefore, should drop to avoid the deadweight loss of tax-motivated consumption.[227]

These results do not hold for risky returns under the same general assumptions because the "next" dollar is not the $(n + 1)^{st}$ but some much lower-income dollar, together with all the others that may be earned up to $(n + 1)$, when the taxpayer makes the investment decision. That is, the more typical setting involves a relatively low-means taxpayer faced with a decision that either will produce a large dollar outcome or nothing (perhaps a loss). The incentives are very different because the low-income level at the time of decision means that income effects are likely to dominate substitution effects.[228] As explained above, the income effect of the imposition of a tax is to engage in more of the income-producing activity (or to consume less). For example, consider an investor facing an investment opportunity that will have payouts of either $10 million or $0 in the next period on an investment of $300,000 in the current period (that is, income of $9,700,000 or a loss of $300,000 in period two). If the government's overall revenue target is met with a flat rate of 30% on all income in all sectors, then the investor's after-tax income (under a full loss offset regime) in the next period is either:

$(1 - t)$*$9,700,000, equal to $6,790,000,

or

$-t$*($300,000), equal to a loss of $210,000.

It seems unlikely that the prospect of a high tax liability would deter the investor given the large overall value of the win (in utility terms) relative to the investor's income at the time the decision is made (zero or close to zero), and given the alternative of much lower income if a different activity is pursued.

The question is whether an increase in taxes on high-dollar wins changes the analysis. If, as optimal tax theory suggests, loss aversion holds, a relatively low-income investor is unlikely to be deterred by a higher tax on a win because the after-tax size of the gain would seem to remain large in utility terms, even for high tax rates, given the investor's low income at the time of the investment. If, for example, the government adopted a tax on high-income successful bets (say, on risky income in excess of $5,000,000) of 50% and made no other changes, the win payoff would become $5,850,000 and the loss payoff would be unchanged.[229] Assuming the $5.9 million payoff remains attractive, the tax does not change investor behavior, and government revenue increases (in the case of the win) by almost $1 million. The increased tax revenue comes at ze-

227. *See, e.g.*, Lawrence H. Summers, *Capital Taxation and Accumulation in a Life Cycle Growth Model*, 71 AM. ECON. REV. 533 (1981) (generally discussing the efficiency losses from taxing returns to savings).

228. Stern, *supra* note 225.

229. The win payoff is (1 – 0.3)*$5,000,000 + (1 – 0.5)*$4,700,000.

ro efficiency cost and can be used to reduce rates in the noninnovation sector.

IV. FULL LOSS OFFSETS: *EX ANTE* AND *EX POST*

To this point, I have argued that tax reform for the innovation sector should seek to remove the extra tax distortions that apply to innovative activity rather than seek to promote innovation directly. The principal policy recommendation that results from the argument is the allowance of full loss offsets—negative taxes for taxpayers with negative incomes. A second recommendation would be elevated rates on high income from innovative activity because the probabilistic nature of returns to innovation means the substitution effects of adopting such rates are minor.

There is, however, a second basis for adopting full loss offsets that is independent of the argument that policy-makers' efforts in the tax area ought to be directed to minimizing tax distortions, though the two arguments dovetail with each other. One of the live debates in innovation policy is whether incentives should take the form of rewards for successful innovation or, instead, should promote innovation up-front, before it is known whether an innovator's efforts will prove successful.[230] *Ex ante* incentives include principally government grants and tax deductions or credits, while *ex post* incentives include principally patents, prizes, and reduced taxes for income from innovative activity.[231] The question, as framed in the tax literature, is partly one of information asymmetry and partly one of capital constraints. Where private actors have a better idea than the government of the costs of innovation an *ex post* incentive system is better because private actors will be better able to determine how much effort to put into an innovation, and the government will not waste resources funding activities with little chance of success.[232] Examples may include consumer-software innovation or other innovations heavily dependent on the use of existing trade secrets held by the innovator. Where the government knows the value better, or where innovators are subject to capital constraints or other frictions that make it difficult for them to finance or pursue a worthwhile project, an *ex ante* regime may prove superior.[233] Examples include capital-intensive industries such as space exploration and socially desirable innovations, such as drug therapies for fatal disease.

230. *See, e.g.*, Duffy, *supra* note 2, at 440–41 (noting the importance of the question of timing of patent issuance in determining an optimal patent regime); Hemel & Ouellette, *supra* note 5, at 333.

231. Hemel & Ouellette, *supra* note 5, at 333.

232. *See* Rachel Glennerster et al., *Creating Markets for Vaccines*, INNOVATIONS 67, 71 (2006), http://www.povertyactionlab.org/sites/default/files/documents/Creating%20Markets%20for%20Vaccines%20Winter%202006_0.pdf.

233. Hemel & Ouellette, *supra* note 5, at 308.

Recent work in property theory by scholars writing in the "new institutional economics" ("NIE") literature provides another reason to favor *ex ante* incentives that is consistent with the efficiency argument offered here. The NIE theorists generally focus on the institutional backdrop to economic activity, investigating the extent to which conventional economic theory presupposes the existence of institutions, be they formal or informal, that themselves cannot be explained in terms of conventional theory.[234] The institution of property has received substantial attention of late. In particular, NIE scholars have challenged the idea that property is properly viewed as an amalgam of more or less arbitrarily identified entitlements, or as a "bundle of sticks" under the common metaphor.[235] The bundle metaphor embodies the idea that the various entitlements associated with traditional ownership, such as the right to occupy, the right to exclude, the right to sell at a given price, and so forth, have no independent or structural significance.[236] Rather, they are discrete and independent of each other, and they can be individually assembled and exchanged on the basis of standard principles of welfare maximization, including, in particular, the idea that markets maximize gains from trade. In short, on the standard view, an item of property is not more than the sum of the entitlements held in the item.[237] The NIE theorists have challenged this understanding, arguing that the property concept has a coherent structure that can be understood, at least partly, in terms of information costs.[238] Property, on this view, is a collection of entitlements that has a specific structure designed to minimize the cost of transmitting information about uses among persons who have no, or only incidental, relationships with each other as regards the thing.[239] Under this view, certain features of property characteristically go together, such as the rights to use and to exclude, the reason being that lumping them in this way becomes efficient once the cost of acquiring information about entitlements is factored in.[240] There are at least two dimensions along which information costs affect the structure of property.

The first dimension sees property as a solution to the problem of coordinating action among largely unconnected individuals, each of whom could, in principle, lay claim to an object for one or more uses.[241] A set of baseline entitlement rules that allocated to a purchaser or otherwise rightful taker only the particular use or uses the purchaser/taker had

234. John N. Drobak, *Law & the New Institutional Economics*, 26 WASH. U. J. L. & POL'Y 1, 1 (2008) ("The new institutionalists believe that economic performance cannot be understood without paying attention to institutions.").

235. Stephen R. Munzer, *Property and Disagreement*, *in* PHILOSOPHICAL FOUNDATIONS OF PROPERTY LAW 289, 289, 291–92 (James Penner & Henry E. Smith eds., 2013).

236. *Id.* at 294.

237. Jane B. Baron, *Rescuing the Bundle-of-Rights Metaphor in Property Law*, 82 U. CIN. L. REV. 57, 58–59 (2014).

238. Henry E. Smith, *Property as the Law of Things*, 125 HARV. L. REV. 1691, 1691 (2012).

239. *Id.*

240. *Id.*

241. *Id.* at 1711.

in mind would create unmanageable complexity with respect to unknown third parties. Upon encountering an object apparently subject to some entitlement, the question would remain whether other entitlements in the object remained available and, if not, who owned them. Conversely, encountering an apparently unburdened object would provide no guarantee that it was not burdened with entitlements for some unspecified uses. Communicating which entitlements are "taken" and by whom and which are "available" under such a regime is prohibitively costly. The property concept, understood as a generalized right to essentially all entitlements in the item (subject to some salient limitations), manages the cost by adopting a simple set of on-off rules that bundles all or most uses in a single "owner."[242] Further allocations of the entitlements to the object can then be negotiated on a case-by-case basis between the property owner and potential users of some of the entitlements in the property who are known to each other.[243] For example, the owner of a farm can lease some of the property to another farmer or to a company that wishes to put a cell-phone tower on the land. The owner can grant mineral rights, grazing rights, air rights, and so forth.

The second dimension of information costs concerns the relationship between the elements of a set of entitlements that the owner anticipates using for various purposes and the potentially larger set of entitlements that the owner actually does use.[244] One of the features of innovative activity is that it is not merely risky, but inherently uncertain. Risk differs from uncertainty in that risk expresses the idea that the likelihood of an outcome can be known *ex ante* even if the outcome itself cannot; whereas, in the case of uncertainty, not even the likelihood of an outcome can be known.[245] For example, the likelihood of randomly drawing a green ball out of a container that is known to have an equal number of randomly mixed red and green balls is one-half. Correlatively, the risk of failing to draw a green ball is also one-half. If, however, the contents of the bag are not known, or only incompletely known, it is not possible to specify the probability of a given outcome or, consequently, of the risk that a given outcome will not materialize; the outcome is simply uncertain.

When an innovator embarks on an effort to invent or discover, the inventor faces uncertainty, not just risk. It generally is not possible to specify the probability of success because we do not know what the range of possible outcomes is beforehand, and we do not know how likely any particular outcome is vis-à-vis other possible outcomes. For example, an effort to invent a particular kind of medication may require more or fewer of the entitlements in various items of property and, perhaps more importantly, new, unanticipated uses for the property may emerge in the

242. *Id.* at 1705–06.
243. *Id.* at 1704–05.
244. *Id.* at 1712.
245. *See* FRANK H. KNIGHT, RISK, UNCERTAINTY, AND PROFIT 20 (1921).

course of innovative efforts. Moreover, because the outcome depends on the innovator's choices and effort, the idea of calculating risk is, to some extent, meaningless—there is no risk profile available for determining how hard or effectively an innovator will work.[246] Indeed, an enduring feature of the history of invention has been that much of what has emerged out of innovative efforts has been found accidentally or unexpectedly. The reason for this, of course, is that by definition the innovator is working in an unknown area on the basis of reasonable, but fallible, beliefs that efforts may bear fruit, without any knowledge of exactly how or what will emerge from the innovative effort. In working with things to invent, the entitlement holder discovers and makes use of features of the things that were not previously known to be of use (or not of use).

This feature of innovation also helps to explain the allocation of entitlement rights in the form of traditional property. Specifically, a second reason entitlements come bundled together, rather than in a more granular specification, is that there are efficiency benefits to assigning the residual entitlement to the person who controls the object.[247] The residual is simply the right to all benefits once all secondary stakeholders have been paid—creditors, customers, employees, and so forth.[248] An alternative would be to hold all nonexpressly allocated entitlements in common. Put otherwise, private ownership would consist only of more narrowly-drawn rights than one finds in the standard property bundle, with the balance available for others' use (either by purchase from a commons or by some other method of claiming them). Because the possibility of some sort of innovation always looms, and because the particular features of a thing that would be employed to pursue an innovation are always somewhat uncertain, the allocation of the residual to the person who, as a first cut, has control over the thing favors innovation: innovations result from use, and the person in possession of the thing is the one most likely to use it.

The significance of the point for tax policy is that the logic bears directly on the choice of whether to provide tax incentives *ex ante* (in the form of loss offsets or, perhaps, refundable tax credits) or *ex post.* An *ex ante* incentive encourages activity that may fail to result in the innovation sought but could lead to other valuable innovation.[249] An *ex post* incentive offers the prospect of rewards solely for sufficiently targeted innovations that are successful; it therefore reduces the impetus to innovate.[250] The efficiency question on *ex ante* versus *ex post* tax incentives, therefore, turns on whether there is a net benefit from giving a tax incentive to the marginal project, defined as the project that is undertaken solely because the tax benefit is categorically available rather than available only

246. YORAM BARZEL, ECONOMIC ANALYSIS OF PROPERTY RIGHTS 4–5 (2d. ed. 1997).

247. *See, e.g.*, Henry E. Smith, *Institutions and Indirectness in Intellectual Property*, 157 U. PA. L. REV. 2083, 2099–2100 (2009).

248. *Id.* at 2092–93.

249. *See id.* at 2118.

250. *Id.*

for targeted innovation that is successful. Marginal projects so defined have three possible outcomes: they result in a targeted innovation, they result in an incidental innovation, or they fail. Thus, the question is whether the benefits from the first two categories are greater than the costs from all three.

It is difficult, if not impossible, to determine the answer to that question empirically, but the same logic that supports giving the residual to the purchaser of a specific, significant entitlement in an object—that is, to making the basic entitlement unit be fee-simple ownership rather than a limited subset of uses—would seem to support a tax incentive regime that applies to all attempts at innovation. One reason is that the likelihood of a successful innovation is, to some extent, endogenous to the incentives that further it: increasing the after-tax expected value of innovative activity makes potential innovators work harder to succeed, and if they work harder, they are more likely to succeed.[251] This feature of the innovation sector is to be contrasted with more-passive stochastic sectors, such as investing in public-traded companies, in which, once a decision to invest is made, the outcome may be uncertain (or risky), but the outcome is not dependent on the investor's attitude about the decision nor on the investor's subsequent efforts made in connection with the decision.[252]

A second reason relates to the value of giving the residual to the holder of property. Giving tax benefits *ex ante* simply completes the property owner's entitlement to the residual—it makes the same tax treatment apply to the residual that applies to the anticipated uses of the property, namely a deduction for costs and taxation of the benefits.

V. CONCLUSION

The temptation to look to tax instruments as a way to remedy perceived market failures in the innovation sector is understandable, but, as I have argued, one that should be mostly resisted. There may be a place for targeted tax incentives to innovation, but a better policy would look first to instruments that correct the market failure in the innovation sector because only these instruments rectify the market failure itself. By contrast, tax solutions may exacerbate the structural market failure by bringing more of the same activity that is subject to market failure to the innovation sector, in addition to introducing new distortions in other sectors.

The observations do not imply that policy-makers should shy from large-scale tax reform for innovation and for risky enterprises generally. Loss limitations in the innovation sector likely have particularly damaging effects on innovation, as a number of scholars have argued.[253] The ef-

251. *See* Henrekson & Sanandaji, *supra* note 134, at 175 (discussing complementarity of capital returns and effort).

252. *Id.*

253. *See supra* Part III.

fects have long been understood to include an incentive not to innovate and an incentive to reduce income volatility through business combinations that, in pretax terms, may be undesirable. To these distortions I have added a third—the incentive to increase risk-taking, as the *ex ante* cost of greater risk-taking under an income tax without loss offsets can actually be lower than that of the less risk-taking, due to the lognormal nature of returns to innovation.

Accordingly, substantial efficiency gains across at least three margins can result from liberalizing the rules on loss offsets and adjusting rates upward to make up for otherwise forgone revenue: less tax-motivated departure from the sector, fewer tax-motivated business combinations, and less untoward risk-taking. Moreover, a further efficiency gain likely is available (as compared to that under flat rates with loss offsets or refundable credits) if policy-makers additionally adopt high marginal rates for innovations that turn out to have big payoffs. This observation is motivated by the standard result in optimal tax theory that a high rate on a large, discrete "win" from a bet is unlikely to discourage the bet. In short, as regards *ex ante* versus *ex post*, it makes more sense to deliver tax benefits up-front, in the form of the reduction in tax disincentives to innovation, and to recoup them at the back end through higher rates, than it does to promise benefits only for successful innovation, as in the case of "patent boxes" and similar regimes.

Finally, and related to the last point, the desirability of *ex ante* rather than *ex post* incentives for innovation is consonant with recent scholarship that explores the nature of property. A tax regime that delivers benefits without regard to the particular innovation that actually results from innovative activity parallels the award of the residual interest in property to its holder. In both cases, the desirability of the incentive is related to uncertainty about the benefits that redound to innovative activity. By giving the benefit in both cases to the holder—ownership of the innovation whether or not the innovation was expected, and a realizable deduction for tax purposes regardless of what results—the rules properly align incentives.