council on economic policies

Intellectual Property Boxes and the Paradox of Price Discrimination

Ben Klemens

May 2017

CEP Working Paper 2017/3

About the Author

Ben Klemens is an Australian economist and author. He works for the US Treasury and was previously a nonresident fellow within the Brookings Institution's Center on Social and Economic Dynamics. Ben holds a PhD in Social Sciences from California Institute of Technology (Caltech). He has also worked on the policy aspects of computing, and in particular the issue of software patents.

The CEP Working Paper Series is published by the Council on Economic Policies (CEP) — an international economic policy think tank for sustainability focused on fiscal, monetary, and trade policy. Publications in this series are subject to a prepublication peer review to ensure analytical quality. The views expressed in the papers are those of the authors and do not necessarily represent those of CEP or its board. Working papers describe research in progress by the authors and are published to elicit comments and further debate. For additional information about this series or to submit a paper please contact Council on Economic Policies, Seefeldstrasse 60, 8008 Zurich, Switzerland, Phone: +41 44 252 3300, www.cepweb.org – info@cepweb.org

Abstract

This paper considers the methods by which some existing laws and proposals offer different tax rates to different types of capital, a scheme variously known as a *patent box*, *innovation box*, or *intellectual property box* (IP box). It presents a model of international tax competition—what tax experts call a *race to the bottom* and competition experts call *Bertrand competition*—with some capital fixed and some easily moved across borders. The model finds that the highest expected tax revenue from mobile IP for a country hosting a large amount of fixed, non-IP capital comes from assigning a single tax rate to all types of capital—that is, from not implementing an IP box. In the context of Bertrand competition, firms optimize revenue when not engaging in price discrimination across types of customers. As a research and development (R&D) credit, several examples show that the IP box is more easily manipulated than a traditional credit on R&D expenses.

Acknowledgements and Disclaimer

Thanks to Lexi Shankster, Susan Chalmers, Annette Deckers, Harry Grubert, Neviana Petkova, William Strang, and Brett York for discussion, suggestions, and critique. Special thanks to Alex Yuskavage for in-depth discussion and some contributions to the text.

Contents

1 Introduction 1

2 Background 4

2.1 How IP revenue crosses borders 4

2.2 Spillovers and returns to invention 6

3 What are we boxing up? 8

3.1 Implementations in law and proposals 8

3.2 Tax incentives 10

3.3 Other R&D incentives 12

4 International Tax Competition 14

4.1 Continuous movement of capital 18

4.2 More than two countries 22

5 Accounting manipulations and back-end subsidies 23

5.1 Shifting expenses to profitable firms 24

5.2 Allocating ambiguous expenses 26

5.3 Inflating R&E costs 27

5.4 Deflating ordinary costs 28

5.5 Choice of invention 28

5.6 A novel method of profiting from patents 29

6 Conclusion 31

Appendix: proofs 33

References 45

1 Introduction

Twelve countries and one Swiss canton have what is variously referred to as an *Innovation Box*, a *patent box*, a *license box*, or an *IP Box*, all of which provide a discounted tax rate for revenue derived from the use of certain intellectual properties. In the 2013–2014 Congressional session, Congressperson Schwartz (D–PA) introduced a bill proposing a patent box that would tax income associated with patents at a 10% rate, rather than the typical 35% corporate rate.[1] In mid-2015, Congresspersons Charles Boustany (R–LA) and Richard Neal (D–MA) proposed a comparable bill that would tax income derived from research and experimentation (R&E) at the lower 10% rate.[2] As of this writing, there are no IP box proposals in Congress, but new proposals may appear in the future, and may use these bills as an initial template.

There are typically two arguments made in favor of IP boxes. The first is that IP boxes are able to keep highly mobile forms of capital within the country. For example, part of the stated intent of both the Schwartz bill and the Boustany-Neal proposal is "to encourage domestication of intangible property". The second is that the lower tax rate helps spur new research and development, as per the title of the Schwartz bill, "The Innovation Promotion Act of 2015".

Others advocate IP boxes as a simple tax break. From Merrill et al. [2012]: "The general objective [of the European IP boxes] is to reduce significantly the corporate tax rate on income from qualifying IP, for example to a nominal rate of 5 to 15 percent, with effective tax rates typically even lower."

This paper divides into three parts, evaluating these arguments from several angles. The first part (Sections 2 and 3) provides background on the IP and tax landscape. Section 2 covers previous research on IP boxes, spillovers, IP transfers, and existing R&D credits, and Section 3 covers the surprisingly wide range of IP box arrangements and underlying definitions.

The second part, in Section 4, examines interjurisdictional competition for intellectual property through tax rates. We consider the baseline case where tax rates on mobile capital also apply to fixed capital, putting a natural brake on how low a

[1] https://www.congress.gov/bill/113th-congress/house-bill/2605/text

[2] http://waysandmeans.house.gov/wp-content/uploads/2015/07/Innovation-Box-2015-Bill-Text.pdf

jurisdiction will be willing to lower tax rates to attract mobile capital; and an IP box alternative where there is a separate rate for profits associated with mobile capital. In one specification considered here, the “race to the bottom” loses any lower limit and the tax rates on mobile capital go to zero. The model is asymmetric, and an IP box implemented by a larger country causes a fall in expected tax rates that an IP box in a small country would not. In another specification, revenue remains positive but unambiguously falls, in line with empirical work such as Griffith et al. [2014], who find that “[p]atent Boxes are likely to attract patent income, but our estimates suggest they will also lead to substantial falls in tax revenues.”

Some tax competition models are primarily aimed at balancing public amenities with tax burdens (such as Wilson [1986] or Zodrow and Mieszkowski [1986]), but these may not be a good fit to describe tax competition over the location of intellectual property licenses, which require no more infrastructure than a mailbox. This model therefore focuses on the question of setting rates given capital in different classes of mobility.

The model and proofs apply equally well to price competition between firms, with one set of customers who are too loyal or locked-in to change given the feasible range of prices, and one set of mobile customers who are price-sensitive. For example, a cable company may have some areas where it is a monopolist and some where it is competing for customers, but for reputational or regulatory purposes may be reluctant to set different prices in different regions. A service with a monthly subscription fee that is auto-deducted from customers’ credit cards will have some customers who don’t pay enough attention to notice a price change, and some who will react to a price change immediately.[3] Some services “lock in” customers to a specific platform; when prices change some users will be able to extricate themselves and others will not. Other game-theoretic treatments of price competition have a homogeneous set of buyers who differ only in willingness to pay [Baye and Morgan, 1999, Kaplan and Wettstein, 2000]; the models here differ because they involve groups of agents who make decisions in fundamentally different manners.

[3] In October of 2015, Netflix, a company that bills customers via a monthly credit card deduction, saw its subscriber count rise by 880,000 over the previous quarter, far below projections of 1.25 million. It attributed the difference to banks replacing older credit cards with new chip-enabled cards and in the process cancelling the previous card number. Evidently, a projected 370,000 subscribers had no problem paying the monthly fee, but didn’t think it worth the effort to re-establish the payment when prompted. See http://www.popularmechanics.com/technology/apps/a17790/netflix-credit-card-chip/.

One might expect the ability to better discriminate among customers or taxpayers would lead to higher revenue. Paradoxically,[4] the model shows that more refined price discrimination can lead to a collapse in revenue that would be impossible when there is no price discrimination.

The proofs are primarily about two players, but many of the key results generalize to three or more players; see Proposition 9.

The IP box is a *back-end subsidy*, meaning that tax benefits are based on profits after sales to customers are completed, as opposed to a *front-end subsidy* on the inputs to the production pipeline. The intent of attracting profits associated with mobile IP demands this structure, but the third part of this paper, Section 5, gives some numeric examples of how a back-end subsidy can be gamed to reduce a firm's taxes without engaging in the R&D the subsidy is intended to induce.

In total, this paper finds little benefit to an IP box. On the international front, intellectual property is uniquely suited to being transferred to the lowest-tax jurisdiction, setting the stage for a race to the bottom game, and an IP box does nothing to mitigate the race, and in fact may remove any natural lower limit to tax rates. In comparison to existing R&E credits which are in direct proportion to R&E spending, some of the characteristics which help a firm take advantage of an IP box have little to do with R&E-related investment. The IP box promises a reduction in average corporate tax rates, but with distortions that are not associated with simpler and more direct means of cutting corporate tax rates.

This paper aims to discuss IP boxes in as general terms as possible, but there are a few ways in which this paper is oriented toward the United States. The first IP box was the Irish box from the 1970s, when Ireland had no tech sector to speak of, and passed this tax incentive in the hopes of attracting or developing one. Conversely, the United States in the present day already has vibrant research and development activity in its borders. The paper is largely theoretical because empirical cross-country comparisons based on a good number of small countries building a tech infrastructure (such as Bradley et al. [2015]) may have limited application to the case of a large country that already has extensive infrastructure. Section 4 presents a model of two countries with different levels of extant capital, and the United States is clearly a high-capital country. The first model specification pre-

[4] Following standard usage in Economics (Edgeworth Paradox, Lucas Paradox, Giffen Paradox, . . .), the word *paradoxically* is used to decribe something initially counterintuitive but readily explained by a careful analysis.

dicts a much more severe effect on tax rates from the U.S. adopting an IP box than from adoption by smaller countries.

2 Background

As IP boxes are a relatively new tax tool, there is little research available on the topic. Evers et al. [2014] provide an excellent overview of various IP box regimes, and raise some concerns regarding implementation. Gravelle [2016a] provides estimates of how an IP box system in the United States might behave.

There is more literature on the underlying motivations behind IP boxes. Internationally, the motivation is in preventing the transfer of ownership of intellectual property to tax havens (Section 2.1). Domestically, the motivation is in spillovers from privately-held intellectual property (Section 2.2).

2.1 How IP revenue crosses borders

An intellectual property right does not typically, in and of itself, generate revenue. A copyrighted work must be printed to a medium that can be purchased or downloaded, a patent must be put into practice on a production line, and a trademark must be prominently affixed to a product. If a branded widget makes a $1 profit, how should the profit be allocated to the physical production and the IP? One might say that without a patent, trademark, or other right to exclude, the widget would produce zero profit, and so all profits are allocable to the IP. Klemens [2015] discusses how this one-drop rule fails in the context of learning by doing and other market imperfections.

But the one-drop rule is preferred by companies with subsidiaries in multiple countries because the mobility of IP makes it easy to allocate profits to subsidiaries in lower-tax jurisdictions.

Consider an independent IP holding company who buys patents from ABC Corp for some cash outlay, then receives future royalties from ABC Corp for its use of those same patents. In a perfectly competitive market, the holding company should make no more than a reasonable return from investing in a risky asset. If it makes a greater return, then ABC Corp could have gotten a better deal from another buyer. Under such theoretically ideal conditions, with no supranormal returns, the tax consequences of the IP reassignment are limited.

But in practice, surprisingly high returns seem exceptionally common with parent-subsidiary arrangements. For example, Apple, Inc (USA) is a 100% owner of Apple Sales International (ASI, Ireland), and the two corporations have a cost-sharing agreement such that both pay for research and both see a return. Ting [2014, p 45] calculated that "the profits to cost ratios under the cost sharing agreement were 7:1 for Apple Inc and 15:1 for ASI" Given ASI's high return, "... it is highly doubtful if Apple Inc would have entered into the cost sharing arrangement with a third party." One might argue that ASI had astute investors with an eye for the best technological bets, but ASI had zero employees until 2012.

Apple is exceptional for being well known and very profitable, but a litany of companies have implemented comparable cost-sharing schemes with subcorporations in lower-tax countries. Samuel M. Maruca, Director of IRS's Transfer Pricing Operations, Large Business & International Division, stated that "as of May 9, 2013, we estimate that we are currently considering income shifting issues associated with approximately 250 [multinational corporate] taxpayers involving approximately $68 billion in potential adjustments to income."[5]

There is some evidence an area's economy does benefit from having patenting activity near to hand [Jaffe et al., 1993], providing part of the justification for tax competition over research and development facilities. But these local spillovers are with regards to the location of patent inventors, not assigned owners of patents. ter et al. [2015] find some expansion of local research activity when a patent box is implemented. But when a patent is moved across borders for tax purposes, do inventors come along? The authors find the opposite: "These results suggest that the tax advantage linked to the patent box does decrease the probability of moving inventors to the patent box country."[p 23]

Arrangements involving only the movement of IP assignments motivate a model with one class of very mobile capital distinct from other classes, where a country or firm sees tax changes from the movement of capital, but no productivity gains or losses. Infrastructure costs for the host country are negligible, because the holding company may have zero employees stationed there.

[5] Senate Testimony by Samuel Maruca to the Senate's Permanent Subcommittee on Investigations, 21 May 2013. http://www.hsgac.senate.gov/download/?id=35C9F013-0B45-43D7-809A-1737DB468345

2.2 Spillovers and returns to invention

Any freely-made exchange produces a surplus for both parties, and some exchanges have ripple effects across the economy, multiplying the surplus they generate. Fundamental to the arguments that patent- or IP-related activity should be subsidized is the argument that IP generates greater spillovers than ordinary goods.

For example, Atkinson and Andes [2011] use the literature finding large IP-related spillovers, discussed below, as evidence that commercial improvements to technology are under-rewarded by the market, and argue that a subsidy in the form of an IP box could compensate for the positive externalities.

This paper will generally use the term *research* to indicate work that produces new intangibles shared among all, and *development* to be the generation of intangibles to be held as private property. Because research produces a public good, the question of research spillovers is only a question of how large an impact the research proves to have. But spending on research can be tenuous, for exactly the reason that its results are not private.

Senator Feinstein's office, borrowing language from the Information Technology & Innovation Foundation, explains that "while R&D tax incentives encourage more basic research, the Patent Box rewards companies that can capitalize on that research and turn a profit."[6] Does the non-research, privately held result of development also have spillovers to other firms or the general economy?

There is some literature, including Griliches [1958] and Tewksbury et al. [1980], where the authors generate a measure of social benefit and use that measure to find that development generates large social gains. Hall et al. [2009] give an especially extensive review of such papers. Following Solow [1956], Nordhaus [2004] finds that profit by firms that initiated an innovation are only a small fraction of the measured productivity gains those firms induced.[7] Malla et al. [2004] uses yield and quality measures of the Canadian canola crop to measure the benefits

[6] http://www.feinstein.senate.gov/public/index.cfm/putting-america-back-to-work or http://www.feinstein.senate.gov/public/?a=Files.Serve&File_id=C7EFE846-C626-40A5-B9B3-1A8B2C5E589A.

[7] Solow presents a macroeconomic model with one good, produced via a function $F(K, L)$ translating the input capital and labor to some amount of the output good. He defines *any* change to the function $F(K, L)$ as *technological change*. This could include improvements in management, surrounding infrastructure, education levels among potential employees, or a host of other factors beyond what we colloquially think of as technology.

from research and development in canola seeds, finding that "the combined effect of IP [rights] and public incentives has driven the quantity of research beyond the socially optimal level."

Social benefit analyses effectively compare a world with a new development to a world without. For one firm, it may be appropriate to make such a comparison in deciding whether it should invest in a research program. For an industry group, such measures may be useful for input-output modeling.

But across the full economy, to compare gains given that a certain inventor invented a widget to a world without the widget is to posit a "Great Man" theory of invention, where only one certain person or group with a certain special genius could conceive of a new invention. But there is a large literature that finds that nearly simultaneous independent invention is common or even the norm.[8]

With this point in mind, it is difficult to apply these studies tallying the total benefit from an invention to the question of a development subsidy. There is greater ambiguity in comparing a subsidized invention to a world where the invention is developed a year—or even a few weeks—later, or where the development evolves in a more distributed manner such that no single patent dominates the market, or where an alternate field of invention solves the same pressing problem. Without a clear description of how the path of invention would differ with or without a subsidy, the value of a subsidy to development is difficult to gauge.

[8] Klemens [2010] takes issue with the popular scientific hagiographies based on a Great Man theory, and gives a sequence of examples covering both pure research and development of inventions, spanning from early electromagnetic research to the modern cell phone, where the first to invent was closely followed by a second-to-invent. Lemley [2012], noting that the idea of simultaneous invention has occurred to many academics, offers a survey of the extensive literature on patent races and independent invention, giving ample evidence that for the great majority of inventions, if the listed inventor were somehow unable to work, the invention would more likely be delayed for some time than taken from humanity forever.

Cotropia and Lemley [2009] looked at real-world situations where two people were practicing the same invention: patent infringement cases. In a world where great inventors are so far ahead that independent invention is impossible, infringement cases would be about copying, not independent invention. Because deliberate copying incurs triple penalties relative to independent invention, plaintiffs have strong incentive to allege copying if there is any evidence of copying to be had. But the authors found that "Only 10.9% of the complaints studied. . . contained even an allegation that the defendant copied the invention, . . . copying was established in only 1.76% of all cases in our data set."

3 What are we boxing up?

There is immense heterogeneity in what constitutes intellectual property, and on top of this a broad range of possible treatments under tax law. Can a movie or musical composition be innovative, and if so, does it merit special tax treatment?

This section considers the definition of IP and R&D under various existing (mostly non-IP box) tax treatments, sometimes overlapping with other definitions and sometimes not. The last part of this section gives a set of examples of how different tax incentives, including IP boxes and existing incentives under U.S. law, measure revenue from IP and how that revenue is treated.

3.1 Implementations in law and proposals

How do countries handle the subtle problem of determining what is IP worthy of special tax treatment? The definitions are as heterogeneous as the types of IP.

This section covers two salient sets of definitions—those based on the United States Code (USC), and those used by the Organization for Economic Cooperation and Development (OECD)—then gives some overview of how those definitions are (or are not) used in the many tax treatments of IP.

The definitions will broadly fall into two categories: input-based, primarily about research effort expended; and product-based, regarding some piece of intangible property such as a patent or a database.

United States

For historic reasons, the first definition of "intangible property" is in the section on taxation of companies in Puerto Rico (PR). Section 26 USC §936(h)(3)(B) defines the term to include, among other things: patents, inventions, formulæ, designs, patterns, know-how, copyrights, trademarks, franchises, methods, programs, customer lists, or "any similar item, which has substantial value independent of the services of any individual."

The definition of *Qualified property* under the Boustany-Neal proposal (herein the B-N definition) keeps only one portion of the PR definition, but adds films and computer programs. This definition therefore covers "Any patent, invention, formula, process, design, pattern, or know-how; Any motion picture film or video type; Any program designed to cause a computer to perform a desired function."

These definitions are clearly product based, focusing on a set of nouns such as patents, formulæ, and know-how, that are markers of the result of R&D or creative labors.

Both definitions go beyond the classes of intellectual property legally protected in the United States (limited to copyrights, patents, trademarks, trade secrets, technical protection schemes under the Digital Millennium Copyright Act, and data sets obtained from some pharmaceutical tests), or even those legally protected in other jurisdictions (such as databases in the European Union). In the other direction, the B-N definition excludes legally protected IP, such as including movies but not music.

As above, 26 U.S.C. §174 allows the expensing of "research or experimental expenditures." The term is not further defined in the statute. But even this much is notable, because the idea of R&E is primarily expenditure- and input-based, as opposed to the product- and output-based definitions above. It allows for labor that produces outputs, such as obvious-to-try inventions, discoveries regarding the laws of nature, or databases of facts, that are outside of U.S. IP law.[9]

[9] The standard regarding obvious to try research was established in *KSR v Teleflex*, 127 S. Ct. 1727.

In Gottschalk v. Benson, 409 U.S. 63, 175 USPQ 673 [1972], the Supreme Court provides a quick overview: "Quoting from earlier cases, we said: 'A principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.' Le Roy v. Tatham, 14 How. 156, 175 [1852]. 'Phenomena of nature, though just discovered, mental processes, and abstract intellectual concepts are not patentable, as they are the basic tools of scientific and technological work.' 409 U.S., at 67, 175 USPQ at 674. 'While a scientific truth, or the mathematical expression of it, is not patentable invention, a novel and useful structure created with the aid of knowledge of scientific truth may be.' 306 U.S., at 94, 40 USPQ at 202."

Regarding databases, see 17 USC §102(b): "In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work." Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991) clarified that even databases produced via a great deal of effort are out of the scope of copyright law.

OECD Frascati

The OECD's Frascati manual [fra, 2002] provides standardized definitions of components of R&D for purposes of statistics and meta-research. Although not intended for the development of tax incentives, its perspective will be useful in our discussion of tax proposals. In §2.1(64), the manual breaks R&D down into three subcomponents: basic research ("experimental or theoretical work"), applied research ("directed primarily towards a specific practical aim"), and experimental development ("directed to producing new materials, products or devices").

The Frascati manual is clearly input-based, and includes research aimed at producing basic scientific results that may be impossible to put into a product-based definition. The Dutch R&D certification (see below) uses a variant of the Frascati definition.

OECD BEPS

The OECD Base Erosion and Profit Shifting project (BEPS) uses this definition of intangibles: "something which is not a physical asset or a financial asset, which is capable of being owned or controlled for use in commercial activities, and whose use or transfer would be compensated" It is a broad product-based definition effectively covering the entirety of the commercially appropriable "knowledge economy" output, including musical compositions or movies.

3.2 Tax incentives

Broadly summarized, IP boxes intend to offer a lower tax rate to a portion of a firm's revenue which can be associated with R&E, so each IP box must specify

1. what portion of the firm's revenue stream qualifies for the IP box, and
2. how to calculate the discount for that portion of the revenue stream.

Beginning with step (1), a *patent box* is a tax incentive tied only to patents; Belgium, France, and the UK are the only countries with patent boxes [Evers et al., 2014].[10] Other countries have *intellectual property boxes* that provide credits for various different types of IP including patents but often going well beyond. For example, the Spanish IP box includes income from production-oriented IP such as patents, plans, secret formulæ, know-how, including royalty income therefrom, but excludes trademarks, copyrights, and software [Murillo, 2015 Basel Congress].

One method of apportionment for step (1) is a *one-drop rule*: if an IP input was used in any way over the course of production, then all profits may be attributed to that IP. Because one can expect that any firm has some sort of qualifying IP (by whatever definition), one could readily argue that every product and service is (following the B-N proposal) "derived from the sale, lease, license, or other disposition of qualified property." In this definition, only profits not tied to a product, such as profitable financial transactions, would be excluded from the IP box.

Here are some examples of interesting implementations. They are intended to give a sense of the breadth of possibilities, but are not a full survey of the space of patent boxes, for which see, e.g., [Evers et al., 2014].

- A qualifying patent for The Netherlands's patent box must be self-developed or "developed for the risk and account of the taxpayer." Further, a profit is only eligible for the 80% tax discount if "the patent right contributes at least 30 per cent to the profit derived..." [Doremaele and Smetsers, 2015 Basel Congress, pp 490–491] There is also a requirement that claimants obtain an R&D certificate for qualifying patents.

- The UK implements the second step above via a paring-down: take the stream of profits derived from the patent, remove the "routine return" (typically about 10%) and "marketing return" to get the patent box eligible revenue [Rabindran and Walsh, 2015 Basel Congress].

- The Hungarian box includes only royalty income, which is given a 50% tax discount, and capital gains from royalty-generating IP, which is not taxed at all [Kolozs and Kőszegi, 2015 Basel Congress].

- A Chinese firm (or a foreign-owned firm in China that meets additional requirements) that has a sufficient percentage of high-tech workers, that has

[10] Supplementary Protection Certificates (SPCs) extend the length of patent protections in the EU for certain patented items, primarily pharmaceuticals, and these patent boxes also cover SPCs.

> a sufficient percentage of its expenses in R&D, that holds some "core IP" to its business, and meets other requirements, is eligible for a 15% tax rate, as opposed to the normal 25% rate. [Stender et al., 2010] The certification is both input-based and product-based. Unlike the other IP boxes, it is the firm, not any one revenue stream, that is deemed to be R&D-intensive.

Adding to the above, we can consider some proposals within the U.S.

Step (1) of the Boustany-Neal formulation uses a one-drop rule that any revenue derived from any of the intellectual property types described in Section 3 would be eligible for the IP box.

Step (2) reduces that profit by the ratio of R&E related costs (as per 26 USC §174) to total costs. Thus, if a company spends 6% of its expenses on R&E, and all of its profits are IP-related, the IP discount would be applied to 6% of profits. The subset of profits that pass both filters pays (in most cases) a 10% rate instead of the typical 35% rate.

This proposal is a hybrid: step 1 is a property-based definition, and step 2 is input-based. If the first step is not interpreted to be so broad as to be vacuous, then it excludes institutions who do research that does not produce exclusive intellectual property. If the first step is taken as vacuous, it reduces to multiplying total revenue by the percent of costs associated with R&E, making the proposal a back-end R&E subsidy. We will compare this form to a front-end R&E subsidy in Section 5.

The Schwartz patent box proposal covers only profits associated with a patent, not the broad list of IP types in the B-N proposal. The input-based R&E-to-ordinary expenses calculation begins by removing a reasonable profit, which is defined as a constant 15%. The remaining profits are then given the reduction from (typically) 35% to 10%.

3.3 Other R&D incentives

The IP boxes are by no means the first efforts to encourage research or development investment. The first is perhaps the patent itself, which under U.S. law is explicitly intended to "promote the Progress of Science and useful Arts",[11] supported by federal outlays on examiners, courts, and enforcement.

Section 174 of Title 26 of the U.S. Code states that "a taxpayer may treat research or experimental expenditures which are paid or incurred by him during the taxable year . . . as expenses which are not chargeable to capital account." That is, rather than depreciating the expense over several years, R&E expenses may be immediately expensed, providing an indirect tax credit on R&E expenses. With an IP box providing a lower IP tax rate, Gravelle [2016a] shows conditions where the IP and the expensing regimes can be revenue equivalent.

If a firm is increasing its research spending over previous years, a portion of the increase can be claimed as a credit. There is also a credit for research in some drug categories.

Localities from nations to cities offer tax discounts to technology-related firms for locating in their area. A 2012 survey from the New York Times tallied $851 million in subsidies 34 U.S. states paid out to tech-sector companies.[12]

As above, firms that have IP that can be transferred abroad often do so, allowing them to shift profits abroad. The revenue stream may eventually be shifted back to the U.S., making the transfer a tax deferral of indefinite duration, or it may remain abroad, obviating the need to pay U.S. taxes entirely. Shay et al. [2016] point out that although it is unlikely that it was intended as such, this is another tax benefit to IP under the existing system, which may encourage firms to expand their IP portfolio.

This paper is intended to discuss IP boxes in as general a setting as possible, and so does not consider their interaction with other incentives. But it is worth bearing in mind that such interactions do exist and may be worthy of future study.

[11] U.S. Const., art. 1, §8, cl. 8.

[12] http://www.nytimes.com/interactive/2012/12/01/us/government-incentives.html#technology

4 International Tax Competition

In all cases, the laws in Section 3.2 define a lower tax rate for profits associated with some type of IP-associated revenue stream. Segregating IP income from ordinary income sets the stage for a tax competition among countries.

As an incentive to redomesticate IP assigned abroad, the perfect IP box would be one that applies a preferential rate only to IP brought from abroad, with zero loss of tax revenue from already-domestic IP royalties that had been taxed at the full rate and now receive a discount. Of course, this ideal is impossible, and there is always some loss of local revenue from even the most targeted policy. Using three estimates by other authors of the elasticity of allocations to tax adjustments, Gravelle [2016b] estimates that for every dollar of revenue lost from offering an IP box discount on already-taxed revenue, re-domestication of IP revenue would yield a 19 cent, 12 cent, or 3 cent gain.

This calculation is based on a single response by IP holders based on a static elasticity. It is therefore not intended to take into account potential after-effects, such as other countries changing their tax rates, or IP holders reassigning their assets back out of the country at a later date. In such a one-shot game, the ideal is a policy that perfectly targets only the mobile IP that could be redomesticated. But the model below shows that such a perfect separation is the worst case for the taxing authority in a strategic game.

The Marrakesh Agreement Establishing the World Trade Organization (also known as the TRIPS agreement) requires that "Members shall accord the treatment provided for in this Agreement to the nationals of other Members."[13] *Nationals* refers both to individuals and corporations, and the treatments provided for includes the holding of patents and trademarks. That is, a country may not impose any sort of nationality requrements on IP holders, an "exit tax", or other fetters to the mobility of intellectual property ownership, leaving tax incentives as one of the few ways in which countries can retain the tax base provided by IP-related sales.

This full mobility facilitates a "race to the bottom" scenario, where tax rates on mobile capital are bid to near zero. This is the outcome predicted by Tiebout-style models of capital mobility (e.g., Tiebout [1956] or Wilson [1986]). And while some

[13] Annex 1C, Art 1, §1(3). https://www.wto.org/english/docs_e/legal_e/27-trips_03_e.htm

economic activities require government infrastructure, which in turn provides a lower feasible bound on tax competition [Plümper et al., 2009], the reassignment of a patent requires only a postal address. So tax rates on IP-related capital have the potential to go even lower than other forms of capital.

To understand the extent to which countries will or will not benefit from this sort of competition, this section presents a model of the choice of tax rates as a game played between two jurisdictions. In the base game, each jurisdiction sets their respective tax rates R_1 and R_2. There is an exogenously fixed upper bound U on R_1 and R_2. In this one-period game, they respectively tax k_1 and k_2 units of revenue associated with fixed capital that is unable to leave the jurisdiction, where $k_1 > k_2 > 0$. The revenue associated with IP capital is the numeraire, fixed at one. It is mobile and will locate itself wherever the tax rate is lower. For completeness, assume that if $R_1 = R_2$ the IP is split between jurisdictions, though this will occur with probability zero in equilibrium.

The fact that all mobile IP moves as a unit makes the game much more tractable, but is not necessarily descriptive, especially given the diversity of IP types from Section 3. Section 4.1 will generalize this game to also include slower-moving capital that follows a smooth transition. As tax rates rise, this first specification of the model follows the norm in Bertrand competition models and has consumers that instantly move to the lower tax jurisdiction, and may be descriptive of firms that already have the legal infrastructure to move IP across borders.

In this game, the mobile capital in all cases is not a decision maker and mechanically responds to R_1 and R_2. Assume also that the quantity of each type of capital is fixed. The only decision is the simultaneous choice of R_1 and R_2 by the players attempting to acquire the tax base of the mobile capital.

Proposition 1. *These strategies describe the unique Nash equilibrium of the base game with fixed and fast-moving capital: jurisdiction two (hosting less revenue from immobile capital) selects a rate by drawing a random value x from a Uniform$[0,1)$ distribution and selecting*

$$R_2 = \frac{Uk_1}{k_1 + 1 - x}.$$

With probability

$$\mu \equiv 1 - \frac{k_2 + 1}{k_1 + 1},$$

jurisdiction one plays U, conceding the mobile capital to player two. Otherwise, it draws a random value x from a Uniform$[0, 1)$ distribution and selects $R_1 = Uk_1/(k_1 + 1 - x)$, just as jurisdiction two does.

The proof that this is a unique Nash equilibrium, as with the proofs of all results to follow, is presented in the appendix.

Note that if $k_1 = k_2$, then $\mu = 0$ and player one's strategy becomes identical to that of player two.

The mixed strategy is a characteristic of Bertrand competition. For example, Blume [2003] presents a Bertrand competition game whose unique Nash equilibrium of the game is a mixed strategy, and in another context, Levitan and Shubik [1972] shows that mixed strategy equilibria may result in the presence of capacity constraints. The mixed strategy can be interpreted as the distribution of outcomes over the course of a large number of similar repeated games, or as a description of what is more or less likely to happen in a one-shot competition between players.

But the best argument for a mixed strategy equilibrium may be that it is the only coherent equilibrium. In a pure strategy equilibrium (R_1, R_2), with $R_1 > R_2$, player one has a strong incentive to switch to $R_2 - \epsilon$, which drives player two to switch to $R_2 - 2\epsilon$, and so on down to a lower bound $Uk_1/(k_1 + 1)$, at which point player one will give up and switch to U—which would lead player two to switch to $U - \epsilon$. In short, any analysis that concludes with a pure strategy can only do so by assuming that one player or the other is unable to change its policies after the initial rates are set. This is true only in the short run.

Returning to the equilibrium of the base game, the jurisdiction with the higher capital level makes no effort to court the mobile capital with probability μ, and simply sets its rate to the maximum U. The strategies of the two jurisdictions are otherwise identical, so with probability $1 - \mu$, both players have an equal chance of winning the mobile capital. Thus, jurisdiction one attracts the mobile capital with probability $(1 - \mu)/2$ and jurisdiction two does so with probability $(1 + \mu)/2$. Returning to the bidding story, in a game between one of the largest economies in the world and a small island nation, μ would approach one and the small island nation would attract the mobile capital with near certainty.

Although k_1 is taken to be fixed in the base game, it is worth considering how outcomes would shift if k_1 were to change.

Proposition 2. *The expected tax rate on mobile capital is an increasing function of the amount of fixed capital k_1 taxed at the same rate.*

This regularity about the tax rate allows the evaluation of the payoff to jurisdiction one.

Corolloary 3. *Jurisdiction one's expected tax revenue from mobile capital only is a nonnegative, increasing function of k_1. As $k_1 \to 0$, revenue goes to zero.*

Corollary 3 shows that, even though the larger jurisdiction is less likely to see any tax revenue from mobile capital as it grows larger, the higher expected tax rate as k_1 rises offsets this. This advises that the larger jurisdiction has an interest in keeping the tax rate associated with mobile capital also associated with as large a fixed capital base as possible.

Choice of game. If the owners of the mobile capital could select the parameters of the game to be played, what would they choose?

By Proposition 2, the expected tax rate on mobile capital falls as k_1 falls.

The IP box is designed to segregate the perfectly mobile IP from the fixed capital, giving profits from each a separate rate. That is, it is intended to push the fixed capital with the same rate $k_1 \to 0$. But in an IP box game where $k_1 \to 0$, expected revenue also goes to zero.

Jurisdiction one is more likely to win the mobile capital as $k_1 \to 0$, but its objective is to maximize expected revenue, not tax base.

We have thus arrived at the first desired comparison of the situation with and without an IP box regime. Without an IP box, taxing profits associated with mobile IP at an ordinary rate, the jurisdictions' hands are tied in their bidding over the profits from mobile capital, and will thus lower the tax rate to an expected value as in Expression 2. With an IP box, the race to the bottom is unfettered, and the expected tax rate approaches zero. Although the larger jurisdiction has a greater chance of having the IP assigned to its jurisdiction with an IP box, it is a Pyrrhic gain, as it draws near-zero taxes from the associated profits.

Note that Proposition 2 (and thus Corollary 3) depends on k_1, not k_2. Jurisdiction two seeks a tax rate just low enough that jurisdiction one is unable to compete, and as above, if k_1 is especially large that limits how low jurisdiction one will go. In terms of changing the game to separate mobile capital from fixed, then, it is jurisdiction one's choice to have or not have an IP box that is most relevant. If jurisdiction one has no IP box and so the similarly-taxed k_1 is large, jurisdiction two has no competitive need to push its IP box rate especially low. If jurisdiction one has an IP box, jurisdiction two will have to compete by pushing its tax rate to zero.

With the segregation of mobile and fixed capital, we can expect that inelastic fixed capital is taxed near U and mobile capital taxed at a rate approaching zero, giving the firm a strong incentive to invent means of associating profits with the mobile capital rate. Such accounting issues are outside the scope of the game described in this section, but will be the subject of Section 5.

4.1 Continuous movement of capital

To this point, the quantity of mobile capital held by country one has been all-or-nothing, but this may not be representative. There is a cost to moving capital, and some capital is semi-fixed, staying where it is until the tax differential is too great to ignore. Such slower-moving capital could be described by a function $k_m(R_1, R_2) \in [0, 1]$ giving the percent of fixed capital allocated to country one, continuous and decreasing in R_1 and increasing in R_2. Country two sees $1 - k_m(R_1, R_2)$ units of slow-moving capital. Although a condition linking the derivatives of $k_m(\cdot, \cdot)$ will be made below, the function need not be symmetric, and allows for a function that expresses a strong preference for one country or the other, thanks to natural or human resources, infrastructure, or other Tiebout-style amenities.

The model in this section uses a fixed linear combination of $(1 - \lambda)$ units of fast-moving capital which instantly jumps to the lower-tax jurisdiction as per the previous model, and λ units of slow-moving capital, where $\lambda \in [0, 1]$, for a total of one unit of mobile capital overall. For example, if country one wins the fast-moving capital, it sees a total capital stock of $k_1 + \lambda k_m(R_1, R_2) + 1 - \lambda$.

A move that is infeasibly expensive in the short run may be inevitable in the long run, which could be accommodated by a λ that trends toward zero as the time frame extends.

There are some proposed standards that would require that tax laws be set so that firms may only benefit from a lowered IP box rate if there is a "nexus of activity" in the country offering the lower rate. That is, a firm that may wish to leave its assets in higher-tax Country A must move some nontrivial amount to Country B to benefit from its lower rate. The nexus of activity scenario is also a story of firms with some combination of fast-moving, slow-moving, and fixed capital, which would be reasonable to describe via some specification of λ, k_1, k_2, and $k_m(\cdot, \cdot)$.

With $k_m(\cdot, \cdot)$ differentiable, the problem is more like a traditional optimization subject to constraints. If the equilibrium point is an interior point, the derivative of the payoff function must be zero there, and the second derivative negative, as with non-game-theoretic optimization problems. However, the optimizations for the two countries are tied together, as country two sees $1 - k_m(\cdot, \cdot)$ in mobile capital, which creates an additional constraint.

We can not solve for the equilibrium without specifying $k_m(\cdot, \cdot)$, but we can state some characteristics of any equilibrium; each proposition in this section will refine what an equilibrium of this game can look like.

With fast-moving capital that moves as a unit as soon as there is the smallest difference in tax rates, the story was that countries have a positive tax rate with some fixed capital in the country, and when fixed capital is removed from the game, the tax rates fall. We find a comparable outcome here:

- As with the model with no slow-moving capital, pure-strategy equilibria with both players selecting the upper bound, both selecting the lower bound, or with both players selecting a tax rate in the interior, are not possible. The pure strategy equilibrium of (U, U) is impossible for the simple reason that either player would prefer to switch to $U - \epsilon$, Lemma 4 shows that the lower bounds will never be part of a pure strategy equilibrium, and Proposition 5 shows that an interior pure strategy equilibrium is impossible.

- Proposition 6 shows that a half-interior strategy is possible, such as (U, R_2) where $R_2 < U$, but this type is only possible when k_1 is large (i.e., $k_1 >$

$k_2 + 1$). This is akin to the solution of the previous model, where country one ceded the race for mobile capital with some likelihood and played U. Unlike the previous model, there may exist conditions where player two cedes the mobile capital, giving an equilibrium of the form (R_1, U), with $R_1 < U$. But as $k_1, k_2 \to 0$, these half-pure equilibria also become impossible.

- The equilibrium of the model with no slow-moving capital was a mixed strategy where both players drew from a distribution that put greater weight on lower tax rates; that is, $p(r)$ is a decreasing function of r. Lemma 7 shows that any mixed-strategy equilibrium over a continuous support will have the same property.

- Proposition 8 echoes the key point of the previous model: In a mixed-strategy equilibrium, the expected payoff from only mobile capital when $k_1, k_2 > 0$ is larger than the expected payoff when $k_1, k_2 = 0$. That is, the game with a unified rate generates more revenue from mobile capital than a game that segregates mobile from fixed capital. However, because players may see some slow-moving capital even when they play U, the expected equilibrium payoff is always greater than zero.

Throughout, assume that $k_m(R_1, R_2)$, representing the capital allocated to country one, is analytic (a smoothness condition meaning that the first, second, third, ... derivatives all exist). The first derivative is negative for player one, positive for player two, and for some proofs we will require symmetric derivatives: at any given point, $-dk_m/dR_1 = dk_m/dR_2$. For one proof, we need concavity (a negative second derivative, which is always necessary at a local maximum).

The proofs are limited to broad statements about the shape of equilibria because they make no further assumptions about the function $k_m(R_1, R_2)$.

Lemma 4. *Assuming $k_m(\cdot, \cdot)$ is analytic, there is always a rate low enough such that the player prefers playing U or close to U over such a low rate.*

Paired with the fact that (U, U) can not be a pure strategy equilibrium, this shows that in any equilibrium, at least one player must be playing an interior rate with nonzero probability. The following propositions therefore seek conditions on what interior equilibria could look like.

Proposition 5. *Given a game with a payoff structure as above, $\lambda < 1$, and $-dk_m(x, y)/dR_1 = dk_m(x, y)/dR_2$ for any (x, y). Then an interior pure strategy equilibrium is impossible.*

Without the assumption of $\lambda < 1$, this Proposition does not prove that $R_1 \neq R_2$. However, if an equilibrium where $R_1 = R_2$ is somehow precluded by the structure of $k(\cdot, \cdot)$, then the value of λ is otherwise irrelevant for the proof, so even when there is no fast-moving capital, an interior pure-strategy equilibrium with $R_1 \neq R_2$ is impossible.

The proof depended on the derivatives of the payoff functions being zero at the optima, but if one optimum is the extremum of U, then the derivative is positive, not zero. When is a pure strategy equilibrium impossible in this case?

Proposition 6. *Given $-dk_m(x, y)/dR_1 = dk_m(x, y)/dR_2$ for any (x, y).*

There can be no pure strategy equilibrium of the form (U, R_2), $R_2 < U$ if $k_1 < k_2+1$.

There can be no pure strategy equilibrium of the form (R_1, U), $R_1 < U$ if $\lambda < 1/2$, or if $k_1, k_2 \approx 0$.

For example, if $k_1, k_2 \approx 0$, there can be no equilibrium of either half-interior form.

These are necessary conditions: the proof made some approximations for the sake of an easier to interpret condition, and the form of $k_m(\cdot, \cdot)$ may still preclude an equilibrium.

For example, the proof regarding the equilibrium of the form (R_1, U) shows that such an equilibrium must satisfy

$$R_1 k_2 > U(k_1 + 1) - (U + R_1)\lambda(1 - k_m(R_1, U)) \tag{1}$$

If $R_1 \to 0$, this can not be satisfied, so $k_m(\cdot, \cdot)$ must be such that the equilibrium point where player one sees its optimum is not be too low.

Because the player strategies are within a finite set ($[0, U] \times [0, U]$), Brouwer's Fixed Point Theorem holds, and we are guaranteed that the game has an equilibrium. If a pure strategy equilibrium is impossible, it must therefore be a mixed strategy equilibrium (and the existence of a pure strategy equilibrium does not preclude a mixed equilibrium). What will a mixed strategy look like in this game with a linear combination of slow-moving and fast-moving capital?

Lemma 7. *In a mixed strategy equilibrium with continuous support, defined by two probability distributions $p_1(r)$ and $p_2(r)$, with $k_m(\cdot, \cdot)$ concave, both distributions are monotonically decreasing in r.*

Proposition 8. *In a mixed strategy equilibrium with continuous support, with $k_m(\cdot, \cdot)$ concave, the payoff from mobile capital only when $k_1, k_2 > 0$ must be greater than the payoff from mobile capital only when $k_1 = k_2 = 0$.*

In the case of a mixed-strategy equilibrium, with the mixture of fixed, fast-, and slow-moving capital we get similar results to the case of only fixed and fast-moving capital: when fixed capital is removed, the expected equilibrium tax rate falls. It may remain greater than zero (depending on the value of $\lambda E_2[k_m(U, \cdot)]$).

For a large country competing with the small island nation, $k_1 > k_2 + 1$, so Proposition 6 does not preclude the larger country ceding mobile capital to the smaller and playing U; this is akin to the point-mass at U in the equilibrium of the model with fast-moving capital only. With $k_1, k_2 \to 0$, this pure strategy of playing U becomes impossible, and we return to the only option being a mixed strategy equilibrium.

4.2 More than two countries

What happens with three or more players? Many of the proofs here are about when an equilibrium is impossible, and if players one and two can not achieve an equilibrium alone, then for the situations here there is no way for then to come to an equilibrium with more players at the table. The results about the possibility of a pure strategy equilibrium where the larger country cedes all mobile capital existing only under certain circumstances and the payoff to a mixed strategy equilibrium declining as $k_1 \to 0$ still hold.

Proposition 9. *Assuming a game where, given the actions of other players, any two players are splitting some remaining slow-moving capital as above, and have some chance of winning the fast-moving capital (which may require selecting one of the players as the low bidder in a proposed pure strategy equilibrium).*

Proposition 5 still holds, guaranteeing no interior pure strategy equilibrium.

Proposition 6 still holds, giving necessary conditions for a half-interior equilibrium.

Proposition 8 still holds, indicating that in a mixed-strategy equilibrium, the expected payoff from only mobile capital when $k_1 > 0$ is larger than the expected payoff from only mobile capital when $k_1 = 0$.

5 Accounting manipulations and back-end subsidies

For the purpose of addressing the international question more easily, the previous section presented a model with three types of clearly delineated capital—revenue from fixed capital could never be reclassified as revenue from mobile. But if there is a separate tax rate for revenue from fixed non-IP capital and revenue from mobile IP, any firm has a strong incentive to characterize as much revenue as possible as being attributable to the lower-rate capital. This section will set aside the international context and consider some examples of how firms within one country can shift revenue into an IP box offered by the country in which it resides.

A back-end subsidy would calculate profits associated with R&D, then apply some lower rate which may itself be determined by choices made by the firm. Broadly, Section 3 discussed two types of allocation of profits to the IP box, the one-drop rule that any IP associated with a revenue stream means all revenue gets the IP box rate, and a proportional allocation depending on the ratio of IP to ordinary expenses. Schemes to reduce tax burden under either allocation rule to will be presented.

This reclassification can be difficult to do with a front-end R&D subsidy, which would ask a firm to report its spending on research or development, then offer a discount on that amount. Under a front-end subsidy, many of the manipulations one could do with a back-end subsidy are impossible:

- A dollar spent by a large firm receives the same credit as a dollar spent by a small firm, so there is no arbitrage to be had in shifting research expenses to different firms. [See the example of Section 5.1.]
- The relative size of the non-R&E budget is irrelevant to the credit, so there is neither benefit to artificially inflating R&E expenses nor artificially understating ordinary expenses. [Section 5.2]

- The loss from research that fails and generates zero profit is lessened by a front-end subsidy but firms must still proceed only if they measure the adjusted risk worth taking. With a back-end subsidy, there are situations where even guaranteed failures will lower overall expenses. [Section 5.3]
- Firms claiming a front-end credit have the straightforward strategy of deducting all expenses. With a back-end subsidy allocated according to the ratio of research expenses to ordinary expenses, firms can avoid taxes by understating ordinary expenses. [Section 5.4]
- A front-end credit is indifferent to ordinary expenses. With a back-end subsidy, there is an incentive to put development effort into products that have large ordinary expenses, so the ordinary expenses can be carried into the IP box. [Section 5.5]
- The purchase of a patent could be structured to qualify for a front-end R&E credit, but the front-end mechanism provides no tax subsidy for using the patent for unproductive lawsuits against practitioners. [Section 5.6]
- Because the R&E credit is input based, not product based, developing a useless patent is a loss-making endeavor, even though the subsidy lessens the loss. There is never a tax incentive to purchase a zombie patent. [Section 5.6]

This remainder of this section presents examples demonstrating how each of these manipulations could work in practice. In each example, the firm receives the subsidy of a lower tax payment, but does no *bona fide* new research, making the back-end subsidy unambiguously worse than a front-end subsidy of the same amount that requires at least some new R&D.

5.1 Shifting expenses to profitable firms

The final IP box rate in a back-end subsidy can change depending on who does the research. Consider two firms: AA Corp, a logistics company, has a 100% profit margin but does no research, while BB Corp has a 20% profit margin on development-heavy goods. The first column of Table 1 summarizes the situation.

In the second column, BB Corp has given operation of its lab to AA Corp, and pays $40 to AA Corp to cover the $40 in lab expenses. In this example, AA Corp

	Baseline	After shifts	No double-counting
AA Corp			
Revenue	100	140	140
Non-R&E expenses	50	50	50
R&E expenses	0	40	40
Pct R&E expenses	0%	44.4%	44.4%
Profits	50	50	50
IP box revenue	0	22.2	22.2
35% tax + 10% IP box tax	17.50	11.94	11.94
BB Corp			
Revenue	60	60	60
Non-R&E expenses	10	10	50
R&E expenses	40	40	0
Pct R&E expenses	80%	80%	0%
Profits	10	10	10
IP box revenue	8	8	0
35% tax + 10% IP box tax	1.5	1.5	3.5

Table 1: In the second column, BB Corp has hired AA Corp to do its R&E. This allows AA Corp to claim a much higher R&E rate.

makes incidental use of BB Corp's products in its work, so by the one-drop rule, all of its profits are now related to its R&E. After the transfer, AA Corp's net revenue is unchanged, but it now has a much higher ratio of R&E expenses to other expenses.

Because hiring a contractor to do R&E counts as an R&E expense, BB Corp's balance sheet and expenditure mix do not change in the second column. The same R&E expense is thus used to adjust the R&E-to-ordinary expense ratio for both AA Corp and BB Corp. In fact, a single R&E expenditure could be chained through as many subcontractors as desired, and each would be able to use that expenditure to put more revenue into the IP box.

Some IP box rules (e.g., France, Belgium) have provisions regarding intermediaries and third-party IP holdings that may prevent double-counting of expenses in this manner. The third column of Table 1 assumes an effective double-counting prevention scheme, so BB Corp counts the $40 it pays to AA Corp as an ordinary expense.

But because AA Corp has saved about 30% on its taxes, it could make a side-payment to BB, or equivalently, allow BB to make a below-cost payment for research expenses. A payment from AA to BB of $3.50 would completely offset all of BB's taxes, still leaving AA $2.06 better off.

This example demonstrates that a dollar spent by a firm with larger profits for whatever reason—larger scale, better marketing, being in a field with higher margins—garners a better tax treatment than identical research done by a firm with lower profits.

5.2 Allocating ambiguous expenses

This example considers the case where the tax rate on income in the IP box is not dependent on the allocation of expenses to R&D (e.g., the UK box); the next example considers the case where the IP box tax rate depends on allocation choices by the firm (e.g., the U.S. proposals).

Evers et al. [2014] notes that if a firm finds a way to expense R&D at normal tax rates and its profits are taxed at IP box rates, then it may be able to achieve a negative tax rate. Table 2 summarizes a firm that initially has an ordinary, non-IP related product. With revenue of $200, expenses of $100, and a tax rate of 35%, the firm nets $65 in profits.

The firm expands to a new IP-related product, with $100 in expenses and $50 in revenue. The middle column of Table 2 shows the situation where the firm allocates its new expenses and revenue to the IP box. At a 10% tax rate, this adds $5 to its tax bill.

But expenses may not be so clearly allocable to the IP-based profits. The expansion may have led to new employees in a larger facility, who may be working on both products at the same time. In the last column, the firm allocates all expenses to the ordinary good. Because it is paying a 10% tax rate on its profits but is deducting 35% for its expenses, its tax payment falls—even relative to its tax payment before it made profits from a new IP-based good.

	Ordinary only	w/IP-based good	Post-reallocation
Non-IP revenue	200	200	200
Non-IP expenses	100	100	100
IP-based revenue	0	100	100
IP-based expenses	0	50	0
IP expenses as ordinary	0	0	50
Taxes	35	40	27.50
Net profits	65	110	122.50

Table 2: If an expense is ambiguously related to IP-based goods (with a 10% tax rate), it is best to claim it as an ordinary expense (at a 35% tax rate).

	Baseline	Do-nothing C_r	Under-report C_o
Revenue	100	100	100
C_d	1	2	1
C_o reported	15.66	15.66	2
C_o unreported	0	0	13.66
35% tax + 10% IP box tax	27.92	26.48	25.86
After-tax profits	55.42	55.85	57.44

Table 3: By increasing its expenses with no revenue effect, or by failing to claim deductions, the firm lowers its tax bill.

5.3 Inflating R&E costs

The last example showed that when the IP box tax rate is fixed, a firm is encouraged to class its expenses as ordinary expenses. This example shows that when the IP box rate depends on IP-related expenses, the firm is encouraged to overstate its IP-related expenses and understate its ordinary expenses.

Let C_o and C_d indicate ordinary and development costs, respectively. Consider a firm whose current R&D budget is such that $C_o = 15.66C_d$, meaning that development is roughly 6% of total costs.[14] Let the ordinary tax rate be $t_o = 35\%$ and the research rate be $t_r = 10\%$, as per the Boustany-Neal proposal. The firm had a hundred units of pre-tax revenue, $P = 100$. Then their after-tax profits are about 55.42. See the baseline column of Table 3 for a summary.

In the next period, the firm doubles its research expenses by hiring a team of researchers to do absolutely nothing, increasing income by exactly $0. After this extra unit of expenditure, $C'_d = 2C_d$, $C'_o = C_o$, and the firm's after tax profits *rise* to ≈ 55.85, an overall gain of $.43$ on its bottom line.

5.4 Deflating ordinary costs

This firm could also get the same result from failing to report non-research expenses. In this example, instead of reporting ordinary costs of $C_o = 15.66$ as above, it reports only costs of $C'_o = 2$. This means that it pays taxes on 13.66 units that it had previously expensed, but because its R&E-to-ordinary ratio shifts so much, its after-tax profits rise to about 57.47, a gain of just over two units. See the last column of Table 3 for a summary.

5.5 Choice of invention

A key problem of the back-end subsidy is that there is no reliable way to determine the extent to which any given component affects the outcome. Theoretical methods such as the Shapley value [Shapley, 1953] that rely on counterfactual valuations would be impossible to administer in tax law. Instead, IP boxes declare some fixed formula to apportion the value of a product to IP versus other inputs. This may be a formula to calculate excess returns as per the UK patent box or a simple percentage, but in no case does it involve an analysis of the technological relation between the invention and the product.

Consider a firm that makes an ordinary, non-IP profit of $100 per year on an ordinary, non-IP related product whose primary expense is aggressive marketing. Its profits are taxed at a 35% tax rate.

[14] Apple, Inc had about this rate of R&D expenses, according to the 10-K it filed with the U.S. Securities and Exchange Commission in 2015.

The firm could develop a minor invention that allows the firm to produce slightly more efficiently (say, saving a few cents per unit). By a one-drop rule, doing so allows it to class the units as IP-related, thus saving 25% on taxes, or $25 per year.

Alternatively, it could put the same effort into developing an entirely new invention. The new invention would have to see an after-tax profit of over $25 before it is more valuable than the tax benefit of the negligible improvement to the existing product. If a formula allocation limited the IP-attached profits to only 50% instead of the 100% of the one-drop rule, the new invention would only see a relative disadvantage of $12.50, but the relative disadvantage would be fully eliminated only when the true added value of the development can be calculated—a potentially impossible feat.

Although the term "fostering innovation" is often used in reference to IP boxes to imply the support of revolutionary new ideas by creative startups,[15] the IP box can push firms to work toward the smallest possible intellectual property claim that will put existing products into the IP box, and lead small firms to push research to large firms that have a larger tax burden to reduce.

5.6 A novel method of profiting from patents

The ideal business model to make use of an IP box might be to start a company that has a mailbox address, pay a lab to develop a patentable invention, wait for third parties to independently implement the patented invention, and sue those third parties. The Boustany-Neal proposal would give a maximal tax discount to this strategy: the proposal takes pains to clarify that compensation for patent infringement is revenue that falls into the IP box, so 100% of our mailbox firm's revenue is IP-related.

Companies that engage in strategies like these are variously called patent assertion entities (PAEs), non-practicing entities (NPEs), or patent trolls.

[15] *Webster's Collegiate Dictionary* of 1977 defines *innvoation* as "the introduction of something new; a new idea, method, or device"; the use of the word to indicate only exceptional or revolutionary change is evidently a recent innovation [Woolf].

As per the derogatory name, there is evidence that patent trolls are a drag on the development of new technology. In a paper entitled "Do Patent Licensing Demands Mean Innovation?" Feldman and Lemley [2015, p 139] answer "Based on our survey results, ex post patent licensing negotiations seem to be almost entirely divorced from innovation." At the height of the patent troll problem, Bessen et al. [2011] found "that NPE lawsuits are associated with half a trillion dollars of lost wealth to defendants from 1990 through 2010. . . . These lawsuits substantially reduce [technology companies'] incentives to innovate." Tucker [2013] did a case study comparing the behavior of firms targeted by a medical imaging PAE to firms not targeted by the PAE. She found that "Relative to firms that were not sued and products that were not covered by the scope of the patent, there was no incremental product innovation in imaging IT by the affected vendors during the period of litigation" Chien [2014] considered who PAEs target. "To the extent patent demands tax innovation, then, they appear to do so regressively, with small companies targeted more as unique defendants, and paying more in time, money and operational impact, relative to their size, than large firms."

Adding to the patent troll problem, a patent box faces a problem with zombie patents. When a patent holder sues another party for infringement of its patent, this initiates an automatic re-examination of the patent by the US Patent and Trademark Office. A careful patent holder therefore never brings a weak, probably invalid patent to trial. Easily invalidated patents granted in good faith by the USPTO are not unheard of: patent-invalidating prior art often surfaces well after a patent is granted, and several rulings by the Supreme Court over the last decade have pulled back the scope of what is patentable, notably in the fields of business methods, software, and medical tests. Many of these effectively worthless patents are technically still in force, because they have not expired, the owner does not involve them in infringement suits where they risk reexamination, and 35 USC §282 states that "a patent shall be presumed valid."

Such zombie patents are an annoyance but not a serious problem under traditional patent law. It is not unusual for firms to have some number of worthless assets, and with reasonable diligence a patent attorney can separate the still-valuable patents from the zombies.

In the tax context, 35 USC §282 and comparable laws in other countries still apply, meaning that productive activity that can be associated with a zombie patent is still eligible for the patent box.[16] There is no method offered in existing IP box proposals for evaluating the validity of a patent. If a patent is invalidated in 2020, there is no provision to undo any tax discounts claimed using the patent in 2018.

The patent box thus gives new life to zombie patents. To give an example, say that a firm makes $100 in profits from a cleverly hedged financial transaction. It pays the full 35% rate on its financial gains. It then buys a patent granted in 2010 on a method for Internet administration of hedged transactions, meaning that its financial transactions are now IP-related as well. Business method patents such as these had been so decisively deemed invalid by *Bilski v Kappos* [561 U.S. 593 (2010)] that the USPTO now has a special examination track (established by Congressional decree) to facilitate invalidating such patents when they are the subject of a reexamination. Nonetheless, writing post-*Bilski*, Rustad [2016, p 927] finds that "An estimated 11,000 patents cover Internet-related business methods."

By buying the patent, a portion of the firm's financial profits move to the patent box, and its tax bill drops from $35 to $15.

6 Conclusion

Due to its ethereal nature, and with the support of trade treaties, intellectual property is possibly the most mobile asset class in existence. Further, the rules that restrict what profits can be allocated to IP are loose, and allow some firms to ascribe large portions of their profits to wherever the IP may be assigned.

The IP box hopes to address this problem by providing an incentive to not assign IP outside of the country. Advocates of IP boxes typically indicate that their intent is also to encourage domestic research and development.

This paper found difficulties with both of these claims. On the international front, it is true that, thinking only one step ahead, a country could gain a short-term benefit from an IP box, but the long run situation after all countries have responded and

[16] As of July 2015, the Schwartz bill even includes profits associated with patents "for which an application is pending before the United States Patent and Trademark Office." [https://www.congress.gov/bill/113th-congress/house-bill/2605/text proposed §200 amendment, paragraph (b)(8)(a)(i), accessed 16 June 2016] But an applicant can file for a patent on anything under the sun, including claims it knows have a minimal chance of passing examination, and can drag on a patent application indefinitely. This clause allows literally anything to be claimed as IP-related, and presses the question of what to do with invalid patent claims much further.

re-responded is better described by the Nash equilibria discussed in the propositions. In these, the mobility of IP makes setting tax rates on IP a race to the bottom. For a larger country like the U.S., expected revenue from taxing mobile IP is increasing in the quantity of non-mobile ordinary capital taxed at the same rate. If the IP boxes achieve their intent and mobile assets are split off to a distinct tax category from nearly all immobile assets, then under a simpler model the expected tax rate on mobile capital falls to zero, and under a more complex model, the expected payoff from mobile capital in a mixed strategy equilibrium unambiguously falls.

Because the primary goal of the IP box is about mobile IP, the subsidy must be keyed to product-based measures such as patents rather than input-based measures of research or development. When considering the effect on domestic firms and their R&D activity, the back-end approach proves to be second-best—imposing distinctions across complex capital types proves to have complications.

The IP box creates numerous undesirable distortions relative to a front-end R&E credit, or direct subsidies to research such as grants by the U.S. National Science Foundation or National Institutes of Health. Firms can manipulate the ratio of research-to-ordinary expenses, lean toward incremental development over new products, eliminate ordinary expenses altogether and become non-practicing entities that make money through the legal system rather than producing products, or purchase business method patents that are useless for anything but reducing the tax rate on financial transactions.

Thus, on the side of corporations seeking to minimize tax burden, an IP box regime requires a number of distortions in its holdings, activities, and balance sheet that would not be necessary under a simpler R&E credit or direct lowering of the corporate tax rate.

On the side of governments seeking to maximize tax revenue, allocation of profits to a patent-holder can currently follow a one-drop rule, but Klemens [2015] suggests that this may be insensible and that an allocation rule that takes into account learning by doing and other non-IP sources of advantage. Although treaties prevent measures to stem the shifting of IP abroad, the allocation of profits to IP depends on accounting laws under each country's control. Tightening those rules and imposing greater scrutiny on transfer pricing statements may be an alternate means of slowing the shifting of profits to offshore havens.

Future directions. The model kept to a general version of the $k_m(\cdot,\cdot)$ function, but this means that the proofs could not present a closed-form construction of the equilibrium as with the model using only fast-moving capital. With a specific definition of this function, one might be able to better specify equilibrium and its characteristics.

What is already here makes empirically testable predictions about the long-run implications of establishing an IP box, notably that revenues will eventually be raced to the bottom. The main barrier in testing this claim is in defining the long run.

As per Section 3, there are diverse types of IP—movies, music, databases, patents, know-how—and different countries have tax laws that cover different sets of IP types. The models presented here assumed all countries are competing for the same IP, but it may be worth considering a world where some countries compete for movies, some for patents, some for both.

Appendix: proofs

Proof for Proposition 1. This section presents a proof that the base game presented in Section 4 has the equilibrium described there, and that this is the unique equilibrium.

Much of the following is symmetric for both players, so use players A and B to stress the fact that the relative capital levels are not relevant.

It is impossible for either player to have a pure strategy equilibrium in this game with continuous strategies. If player A chooses fixed strategy R_A, then player B could choose one of $R_B = U$ or $R_B = R_A - \epsilon$, depending on which has the higher payoff. But (R_A, U) is not an equilibrium because player one would prefer to choose $U - \epsilon$ over R_A, and $(R_A, R_A - \epsilon)$ is not an equilibrium because player one would prefer to play one of $R_A - 2\epsilon$ or U.

If there is any range in A's strategy (R^m, R^M) that A plays with probability zero, then B must play there with probability zero as well, because any $R_B \in (R^m, R^M)$ is dominated by R^M. But if B is playing in this range with probability zero, then A's choice of R^m is dominated by $(R^m + R^M)/2$. Therefore, the support of the equilibrium strategy for both players must span without breaks from some lower bound up to U.

Similarly, the lower bound of both players' strategies must be identical. If the lower bounds for A and B are $R_A^m > R_B^m$, then any choice for B in the range of $[R_B^m, R_A^m)$ is dominated by R_A^m. Therefore, the support for both players must be from some common lower bound up to U.

One player may have a point-mass probability at U, so that if A has probability m of playing U, $P(R_A < U) = 1 - m$. Then there is a discrete difference in odds of winning the mobile capital between playing $R_B = U$ and $R_B = U - \epsilon$, so B will strictly prefer $U - \epsilon$. That is, if A plays a point-mass strategy at U, B will play a strategy over the support $[R_B^m, U)$. That is, it is impossible to have an equilibrium where both players put a point mass on U.

Player A prefers the fixed strategy of U (with payoff Uk_A) over any value of $R_A < Uk_A/(k_A + 1)$, because even if guaranteed the unit of mobile capital, the payoff would be less than Uk_A. Therefore no rate less than $Uk_A/(k_A + 1)$ can be in the support of the equilibrium.

If player A plays U, it must be indifferent between losing the mobile capital with certainty and all other options, so its expected payoff must be Uk_A. This need not be the case with the other player. At the bottom of the support of B's strategy R_B^m, the probability of gaining the mobile capital must be one, so at that point $R_B^m(k_A + 1) = Uk_A$, so $R_B^m = Uk_A/(k_A + 1)$.

One player or the other has a lower bound of its equilibrium support of $Uk_A/(k_A + 1)$. Because $Uk_2/(k_2 + 1) < Uk_1/(k_1 + 1)$, and the support of both sides of the equilibrium must be identical, it must be the case that the player who puts nonzero weight on U is player one, so the support of player one's strategy is $[k_1/(k_1 + 1), U]$ and the support of player two's strategy is $[k_1/(k_1 + 1), U)$. These are the unique supports for any equilibrium.

If player one selects an arbitrary R_1 from the range $[Uk_1/(k_1+1), U]$, it can be written as $R_1 = Uk_1/(k_1+1-x)$ for some $x \in [0,1]$. Player two selects R_2 via some other random process, and write $P(R_2 > R_1)$ for the likelihood that player one attracts the mobile capital. Then the expected payoff to player one is $Uk_1/(k_1+1-x)[k_1 + P(R_2 > R_1)]$. This must equal Uk_1 for all x in the range, meaning $1 - x = P(R_2 > R_1)$, which is true only when two selects $x' \in [0,1]$ at random and sets a rate of $R_2 = Uk_1/(k_1+1-x')$.

Write the odds of player one choosing U as μ. For player two to be indifferent over the same range from $k_1/(k_1+1)$ to U, similar linear equations give the strategy for player one as in the text. Player two will select some value R_2 which can be expressed as $R_2 = Uk_1/(k_1+1-x)$, for some x. It always taxes its fixed capital k_2, with probability μ it wins the unit of mobile capital with certainty, and in the probability $1-\mu$ case that player one competes for the mobile capital, player two's draw of x will be smaller with probability $1-x$. Then player two's overall payoff is

$$R_2[k_2 + \mu + (1-x)(1-\mu)].$$

For this to be an equilibrium, this expression must be constant across all values of x. We can verify that if

$$\mu = 1 - \frac{k_2+1}{k_1+1},$$

player two expects the same payoff for any x in the support of its strategy:

$$\begin{aligned}
& R_2[k_2 + \mu + (1-x)(1-\mu)] \\
= \; & \frac{Uk_1}{k_1+1-x}\left[k_2 + 1 - \frac{k_2+1}{k_1+1} + \frac{(1-x)(k_2+1)}{k_1+1}\right] \\
= \; & \frac{Uk_1(k_2+1)}{k_1+1-x}\left[\frac{k_1+1}{k_1+1} - \frac{1}{k_1+1} + \frac{1-x}{k_1+1}\right] \\
= \; & \frac{Uk_1(k_2+1)}{k_1+1}
\end{aligned}$$

Thus, the given strategies are the only solution of the linear equations that generate indifference across all elements in the unique supports for the equilibrium strategies. □

Proof for Proposition 2. In the probability μ contingency where player one selects U, the tax rate becomes whatever player two chose, based on player two's draw from a Uniform Distribution. The probability of any given draw is simply $P(x) = 1$. With probability $1 - \mu$, both players draw from a Uniform Distribution and the smaller draw becomes the tax rate. The odds that x is the smallest of two draws is described by a Beta(1,2) distribution, $P(x) = 2(1 - x)$.

After the draw, the tax rate is $R = Uk_1/(k_1 + 1 - x)$. For the case of a single draw ($P(x) = 1$), the expected value of R is

$$\int_0^1 \frac{Uk_1}{k_1 + 1 - x} dx = Uk_1 \left(ln(k_1 + 1) - ln(k_1) \right). \tag{2}$$

This is increasing in k_1.

In the case where x is the smaller of two draws, the expected value of R is

$$\int_0^1 \frac{2Uk_1(1 - x)}{k_1 + 1 - x} dx \quad = \quad 2Uk_1 \left(k_1 \ln(k_1) + 1 - k_1 \ln(k_1 + 1) \right)). \tag{3}$$

This is also increasing in k_1.

For any value of k_1, Expression 2 is greater than Expression 3, as one would expect given that the first is based on a single draw and the second on the minimum of two draws. Also, μ is an increasing function of k_1, meaning that Expression 2 is more likely to be selected given a larger value of k_1. Combining these points confirms the claim that the expected rate R is increasing in k_1. □

Proof for Corrolary 3. Jurisdiction one's payoff from taxing mobile (not fixed) capital is zero if it does not win the mobile capital. It competes with probability μ, and if it does compete it wins with probability $1/2$. When it does win, its expected payoff is as per Expression 3. Joining these together, its expected payoff is

$$\frac{k_2 + 1}{k_1 + 1} Uk_1 \left(k_1 \ln(k_1) + 1 - k_1 \ln(k_1 + 1) \right))$$

Both $k_1/(k_1 + 1)$ and $\ln(k_1) - \ln(k_1 + 1)$ are increasing functions of k_1, making the whole an increasing function of k_1, which was to be shown. As $k_1, k_2 \to 0$, this expression goes to zero.

Proof for Lemma 4. Given that $k_m(\cdot, \cdot)$ is analytic, the *Principle of Permanence* applies. The Principle states that if a function is zero over a range of more than measure zero, then it must be zero everywhere. It is possible that $k_m(U, R_2) = 0$ for all R_2, but if $k_m(U - \epsilon, R_2)$ is also zero for all ϵ below some value, then the PoP states that $k_m(R_1, R_2) = 0$ for all values of R_1, R_2—there is no slow-moving capital, and we return to the proofs in the previous section with only fast-moving capital. So there must be a positive payoff at U or $U - \epsilon$.

The payoff at $R_1 = 0$ is zero; by continuity, there must be some choice of small value of R_1 where the player is indifferent between the lower and higher rate (and if the payoff is always zero, the PoP again holds and proves there is no mobile capital). Therefore, there can never be a non-interior equilibrium at the lower bound.

Proof for Proposition 5. Define W to be one if $R_2 > R_1$ and player one wins the fast-moving capital, and zero if $R_1 > R_2$ and player two wins. Let the capital levels be

$$\kappa_1 \equiv k_1 + \lambda k_m(R_1, R_2) + W(1 - \lambda) \tag{4}$$

$$\kappa_2 \equiv k_2 + \lambda(1 - k_m(R_1, R_2)) + (1 - W)(1 - \lambda) \tag{5}$$

Then the payoff functions are $P_1(R_1, R_2) = R_1\kappa_1$ and $P_2(R_1, R_2) = R_2\kappa_2$.

Consider an equilibrium where both players play the same tax rate, $R_1 = R_2$ (possibly U). This can not be an equilibrium because player 1 will want to change its tax rate to $R_1' = R_2 - \epsilon$, and similarly for R_2. By Lemma 4, $R_1 = R_2 = 0$ can not be an equilibrium.

Therefore, any pure strategy equilibrium must have different tax rates.

The derivatives of the payoff functions are zero at their optima:

$$0 = \tfrac{dP}{dR_1} = \kappa_1(R_1) + \lambda R_1 \frac{dk_m}{dR_1}(U, R_2)$$

$$0 = \tfrac{dP}{dR_2} = \kappa_2(R_2) - \lambda R_2 \frac{dk_m}{dR_2}(U, R_2)$$

Note that a change of $\pm\epsilon$ can not change who wins the mobile capital, so W is treated as constant for the derivative.

Solving for dk_m/dR in both cases, and using the symmetry assumption, gives

$$\frac{\kappa_1}{\lambda R_1} = -\frac{dk_m}{dR_1} = \frac{dk_m}{dR_2} = \frac{\kappa_2}{\lambda R_2}.$$

Expanding the κs, we can solve for $k_m(R_1, R_2)$ at any ostensible equilibrium:

$$k_m(R_1, R_2) = \frac{R_1[k_2 + \lambda + (1-W)(1-\lambda)] - R_2[k_1 + W(1-\lambda)]}{\lambda(R_1 + R_2)} \tag{6}$$

Assume that $R_1 > R_2$. Because we will not use the fact that $k_1 > k_2$, this is without loss of generality, and the same proof could be redone with $R_2 > R_1$ (basically swapping $k_m(\cdot,\cdot)$ and $1 - k_m(\cdot,\cdot)$ until the symmetric Inequality 9).

Player 1 sees none of the fast-switching capital, but could if it switches to $R_2 - \epsilon$. The fact that it does not implies:

$$R_1[k_1 + \lambda k_m(R_1, R_2)] > (R_2 - \epsilon)[k_1 + \lambda k_m(R_2 - \epsilon, R_2) + 1 - \lambda]$$

Player 2 could switch to a higher rate of $R_1 - \epsilon$ without losing its lock on the fast-switching capital. That it does not do so means:

$$R_2[k_2 + \lambda(1 - k_m(R_1, R_2)) + 1 - \lambda] > (R_1 - \epsilon)[k_2 + \lambda(1 - k_m(R_1, R_1 - \epsilon)) + 1 - \lambda]$$

Joining the two conditions (and rounding off some irrelevant ϵs):

$$\frac{k_1 + \lambda k_m(R_1, R_2)}{k_1 + \lambda k_m(R_2 - \epsilon, R_2) + 1 - \lambda} > \frac{R_2}{R_1} > \frac{k_2 + \lambda(1 - k_m(R_1, R_1 - \epsilon)) + 1 - \lambda}{k_2 + \lambda(1 - k_m(R_1, R_2)) + 1 - \lambda} \tag{7}$$

Substituting $k_m(\cdot,\cdot)$ from Equation 6 in the appropriate points in Inequality 7, with W equal to zero or one as appropriate, expands to

$$\frac{k_1 + \lambda \frac{R_1(k_2+1) - R_2 k_1}{\lambda(R_1+R_2)}}{k_1 + \lambda \frac{k_2 + \lambda - k_1 - 1 + \lambda}{2\lambda} + 1 - \lambda} > \frac{k_2 + \lambda(1 - \frac{k_2 + 1 - k_1}{2\lambda}) + 1 - \lambda}{k_2 + \lambda(1 - \frac{R_1(k_2+1) - R_2 k_1}{\lambda(R_1+R_2)}) + 1 - \lambda} \tag{8}$$

Let $k_s \equiv k_1 + k_2 + 1$. Inequality 8 simplifies remarkably well, to

$$\frac{\frac{R_1 k_s}{R_1 + R_2}}{\frac{k_s}{2}} > \frac{\frac{k_s}{2}}{\frac{R_2 k_s}{R_1 + R_2}}$$

Taking a cross-product, the equilibrium holds when

$$\left(\frac{R_1}{R_1+R_2}\right)\left(\frac{R_2}{R_1+R_2}\right) > \left(\frac{1}{2}\right)^2 \tag{9}$$

Now let $\Delta \equiv (R_1 - R_2)/(2(R_1 + R_2))$, the distance from $R_1/(R_1 + R_2)$ to a half. Then Inequality 9 becomes

$$\begin{aligned}\left(\frac{1}{2}+\Delta\right)\left(\frac{1}{2}-\Delta\right) &> \left(\frac{1}{2}\right)^2 \\ \frac{1}{4}-\Delta^2 &> \frac{1}{4}\end{aligned}$$

With $R_1 \neq R_2$, the left-hand side is strictly less than 1/4, so this condition can never be satisfied. □

Proof for Proposition 6. Consider the case where at equilibrium $R_1 = U$ and $R_2 < U$. Player two does not want to change its strategy by $\pm\epsilon$, meaning that dP_2/dR_2 is zero. Player one chooses an extreme value because the slope dP_1/dR_1 is still upward at the edge of the space:

$$\begin{aligned}0 < \tfrac{dP}{dR_1} &= \kappa_1(R_1) + \lambda R_1 \frac{dk_m}{dR_1}(U, R_2) \\ 0 = \tfrac{dP}{dR_2} &= \kappa_2(R_2) - \lambda R_2 \frac{dk_m}{dR_2}(U, R_2)\end{aligned} \tag{10}$$

Solving for dk/dR_i for $i = \{1, 2\}$:

$$\begin{aligned}-\frac{dk_m}{dR_1}(U, R_2) &< \frac{\kappa_1}{\lambda U} \\ \frac{dk_m}{dR_2}(U, R_2) &= \frac{\kappa_2}{\lambda R_2}\end{aligned}$$

By the symmetry assumption, the two derivatives are equal, so $R_2\kappa_1 > U\kappa_2$. This expands to

$$\begin{aligned}R_2[k_1 + \lambda k_m(U, R_2)] &> U[k_2 + \lambda(1 - k_m(U, R_2)) + 1 - \lambda] \\ R_2 k_1 &> U(k_2 + 1) - (U + R_2)\lambda k_m(U, R_2)\end{aligned}$$

If $k_1 < k_2 + 1$, this can never be satisfied.

Repeating the procedure with $R_1 < U$, $R_2 = U$, and the derivative conditions reverse, and we need $R_1\kappa_2 > U\kappa_1$, which expands to

$$\begin{aligned} R_1[k_2 + \lambda(1 - k_m(R_1, U))] &> U[k_1 + \lambda k_m(R_1, U) + 1 - \lambda] \\ R_1 k_2 &> U(k_1 + 1) - (U + R_1)\lambda(1 - k_m(R_1, U)) \end{aligned}$$

If $\lambda < 1/2$, or if $k_1, k_2 \to 0$, the condition can never be satisfied. □

Proof for Lemma 7. Let $p_2(\cdot)$ be the equilibrium probability distribution of tax rates for player two, and let $E_2[f(R)]$ indicate the expected value using measure $P_2(\cdot)$: $\int_{\forall R} f(R)P_2(R)dR$.

Both players must be indifferent between all options in the support of the equilibrium, so for player one the payoff

$$R_1 \left[k_1 + (1 - \lambda) \int_{R_1}^{U} p_2(r)dr + \lambda \int_{R_1^{min}}^{U} p_2(r)k_m(R_1, r)dr \right] \tag{11}$$

must be constant for all values of R_1. Let $\kappa_1(R_1)$ indicate the capital level in brackets; note that it is a function only of R_1, as R_2 is integrated out.

For the payoff to be constant, its derivative must be zero:

$$0 = \kappa_1(R_1) + R_1 \left[-(1 - \lambda)p_2(R_1) + \lambda \frac{dE_2[k_m(r, \rho)]}{dr} \right]$$

With $k_m(\cdot, \cdot)$ analytic, the derivative exists so long as $p_2(\cdot)$ has only a finite number of discontinuities. Solving for $p_2(R)$:

$$p_2(r) = \frac{\kappa_1(r)}{r(1 - \lambda)} + \frac{\lambda}{1 - \lambda} \frac{dE_2[k_m(r, \rho)]}{dr} \tag{12}$$

Because $k_1(R_1^a)R_1^a = k_1(R_1^b)R_1^b$, for all a, b, we can write $k_1(R_1) = C/R_1$, where C is a constant based on parameters of the situation but constant for all tax rates, and write $(1 - \lambda)\kappa_1(r)/r$ as $(1 - \lambda)C/r^2$.

With a negative second derivative of $k_m(r,\rho)$, the second derivative of $E_2[k_m(r,\rho)] = \int k_m(r,\rho)p_2(\rho)d\rho$ must also be negative for a given $p_2(r)$.[17] That is, $dE_2[k_m(r,\rho)]/dr$ is smaller for larger r.

Thus, both terms in Equation 12 are monotonically decreasing. □

Proof for Proposition 8. Begin with $k_1 > 0$, then consider a new game with $k_1' < k_1$ (possibly $k_1' = 0$). The elements of the game with $k_1 = 0$ will be denoted with primes (C', $\kappa_1'(U)$, $p_2'(r)$, . . .).

Let $\Delta_p(R_1) \equiv p_2'(R_1) - p_2(R_1)$

Consider the a case where there is a point R_1^t such that $\Delta_p(r) < 0$ for all $r > R_t$, and $\Delta_p(R_1^t) = 0$. Because $\Delta_p(R_1)$ must integrate to zero, this indicates that there is some point R_1^b where the total for Δ_p between R_1^b and R_1^t must be positive.

To simplify notation, let $M(R_1)$ be the total expected mobile capital at R_1, so $\kappa_1(R_1) = k_1 + M(R_1)$. Then, with $R_1^b < R_1^t$ (so $M(R_1^b) > M(R_1^t)$ and $M'(R_1^b) > M'(R_1^t)$):

$$\frac{k_1 + M(R_1^t)}{k_1 + M(R_1^b)} = \frac{R_1^b}{R_1^t} = \frac{M'(R_1^t)}{M'(R_1^b)}.$$

Thus,

$$\frac{M(R_1^t)}{M(R_1^b)} < \frac{M'(R_1^t)}{M'(R_1^b)}. \tag{13}$$

Then

$$\begin{aligned}
\int_{R_1^t}^{U} p_2'(r)dr &= \int_{R_1^t}^{U} p_2(r)dr + \int_{R_1^t}^{U} \Delta_p(r)dr \\
\int_{R_1^b}^{U} p_2'(r)dr &= \int_{R_1^b}^{U} p_2(r)dr + \int_{R_1^t}^{U} \Delta_p(r)dr + \int_{R_1^b}^{R_1^t} \Delta_p(r)dr
\end{aligned}$$

[17] If $k_m(\cdot,\cdot)$ is analytic, then

$$d^2\frac{\int k_m(r,\rho)d\rho}{dr^2} = \int \frac{d^2 k_m(r,\rho)}{dr^2}d\rho.$$

With the integral from R_1^t to U negative, and the integral from R_1^b to U less negative (or zero), we have

$$\frac{\int_{R_1^t}^U p_2'(r)dr}{\int_{R_1^b}^U p_2'(r)dr} < \frac{\int_{R_1^t}^U p_2(r)dr}{\int_{R_1^b}^U p_2(r)dr}$$

For the slow-moving capital, assume the same setup: above some point R_1^t the total Δ_p is less than zero, and there is a breakeven point R_1^b where the total Δ_p is zero. Define

$$\delta^b \equiv \int_{R_1^b}^{R_1^t} \Delta_p(r)k(R_1^t, r)dr > 0.$$

Because $k_m(\cdot, \cdot)$ is increasing in the second argument,

$$\delta^t \equiv \int_{R_1^b}^{R_1^t} \Delta_p(r)k(R_1^t, r)dr + \int_{R_1^t}^{R_1^U} \Delta_p(r)k(R_1^t, r)dr < 0$$

Because the expected value integral is over the entire range of rates, we may need to sum a sequence of such δ^ts in order to cover sets of negative then positive or net-zero Δ_ps.

$$\frac{E_2[k_m'(R_1^t)]}{E_2[k_m'(R_1^b)]} = \frac{E_2[k_m(R_1^t)] + \delta^t}{E_2[k_m(R_1^b)] + \delta^b} < \frac{E_2[k_m(R_1^t)]}{E_2[k_m(R_1^b)]}$$

Combining both slow and fast moving capital, we have a contradiction to Inequality 13. The implication is that there can be no point R_1^t such that the total Δ_p above that point is less than zero.

Therefore, consider the case where there is a point R_1^t such that $\int_r^U \Delta_p(r)dr > 0$. That is, for large values of r, $p_2'(r) > p_2(r)$, which requires for small values of r, $p_2(r) > p_2'(r)$. Then $M(R_1) > M'(R_1)$, as both the fast-moving and slow-moving capital terms put more weight on lower-capital rates and so are unambiguously smaller in the primed case. If Δ_p vacillates between being positive in some parts but negative in others, all while $\int_r^U \Delta_p(r)dr > 0$, the full range can be broken down into paired segments where the upper segment has $\Delta_p > 0$ and the lower $\Delta_p < 0$; in each, the upper segment loses more from $M(R_1)$ than the lower segment gains, for an overall decrease.

One final consideration: the bottom of the support could change, from R_1^m to $R_1^{m'}$. If $R_1^m < R_1^{m'}$, then the statement that the expected value $M(R_1)$ is greater than

$M'(R_1)$ still stands, as the relationship has been proven for that density over $R_1^{m'}$, and the fact that $M(R_1)$ has still more density below $R_1^{m'}$ only bolsters the inquality.

For the case when $R_1^m > R_1^{m'}$ we need to check whether the new density in the range $[R_1^{m'}, R_1^m]$ change this inequality.

For the fast-moving capital, where for a given R_1 the expected capital intake is $(1-\lambda)\int_{R_1}^{U} p_2'(r)dr$, the density below $R_1^{m'}$ for $p_2(r)$ is simply zero, so the integral is one; with some density in $p_2'(\cdot)$ below $R_1^{m'}$, this value unambiguously falls to below one.

For the slow-moving capital, $k_m(R_1, r) \geq k_m(R_1, R_1^m)$ for all $r > R_1^m$, and the reverse for all $r < R_1^m$. Let Δ_p^m the the density of $p_2'(r)$ in $[R_1^{m'}, R_1^m]$, meaning that there is a loss of Δ_p^m in density above R_1^m. Then the expected slow-moving mobile capital is

$$\lambda \int_{R_1^{m'}}^{U} p_2'(r) k_m(R_1, r) dr = \lambda \left[\int_{R_1^m}^{U} p_2(r) k_m(R_1, r) dr - \Delta_p X + \Delta_p Y \right], \tag{14}$$

where X is some value greater than $k_m(R_1, R_1^m)$ representing a weighted average of capital levels above R_1^m and Y is a value representing a weighted average of capital levels below the former minimum. We must have $X < Y$, so $E_2'[k_m(R_1, \rho)] < E_2[k_m(R_1, \rho)]$.

Thus, for this term as well, the lowering of $p_2(r)$ effectively means that player two underbids player one more often, and we see lower expected capital for both the fast-moving and slow-moving capital.

In total, for any given R_1 and any admissible shift from $p_2(r)$ to $p_2'(r)$, we have $M(R_1) > M'(R_1)$.

If we had left k_1 in place, we would have $R_1(k_1 + M(R_1)) > R_1 M'(R_1)$; that is, $C > C'$, or, given that the Cs are constant, $E[C] > E[C']$.

Let $M(U)/U \equiv C^U$. As above, $M(U) > M'(U)$, so $C^U > C'$. For values $r < U$ and $U(k_1 + M(U)) = r(k_1 + M(r)) = C$, eliding k_1 means $rM(r) > UM(U) = C^U$. Therefore, for any R_1,

$$R_1 M(R_1) > C^U > C' = R_1 M'(R_1).$$

Then $E[R_1 M(R_1)] > E[C'] = E[R_1 M'(R_1)]$, thus establishing that the expected value in the game with $k_1 = 0$ has a lower payoff than the expected payoff from only mobile capital in the game with $k_1 > 0$.

The fact that $k_1 > k_2$ was unused, so the result holds symmetrically for player two. $\square$

Proof for Proposition 9. For Proposition 5, consider a proposed pure strategy equilibrium. As per the proof, it is impossible that the two lowest bidders bid the same value, so any pure strategy equilibrium must have two low bidders at different values. The proof holds for those two lowest bidders to show that they are not in an equilibrium.

Proposition 6 applies pairwise to any two bidders.

The proof of Proposition 8 is from the perspective of player one, deciding what to play give player two's probability distributions $p_2(r)$ and $p_2'(r)$. Notably, those distributions must make player one indifferent over the support of its equilibrium. It is effectively two separate proofs: one regarding the fast-moving capital and one regarding the slow-moving capital.

Let $p_F(r)$ be the aggregate likelihood that r is the lowest selected value among all of one's opponents, and let $p_S(r)$ be an aggregate probability such that $E_S[R_1]$ has the desired value for any given R_1. Then the proof goes through with $p_2(r)$ replaced with either $p_F(r)$ or $p_S(r)$ in the appropriate segments of the proof.

References

Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development. OECD Publishing, December 2002. doi: 10.1787/9789264199040-en.

Robert D Atkinson and Scott Andes. Patent boxes: Innovation in tax policy and tax policy for innovation, October 2011. URL http://www.itif.org/files/2011-patent-box-final.pdf.

Michael R Baye and John Morgan. A folk theorem for one-shot bertrand games. *Economics Letters*, 65(1):59–65, 1999. doi: https://doi.org/10.1016/S0165-1765(99)00118-4.

James Bessen, Jennifer Ford, and Michael J Meurer. The private and social costs of patent trolls. *Boston University School of Law Working Paper Series*, 11(45), November 2011.

Andreas Blume. Bertrand without fudge. *Economics Letters*, 78(2):167–168, February 2003. doi: 10.1016/s0165-1765(02)00217-3.

Sebastien Bradley, Estelle P Dauchy, and Leslie A Robinson. Cross-country evidence on the preliminary effects of patent box regimes on patent activity and ownership. (2681433), October 2015. doi: dx.doi.org/10.2139/ssrn.2681433.

Colleen Chien. Startups and patent trolls. *Stanford Technology Law Review*, 17: 461–506, 2014. URL http://str.stanford.edu/pdf/startupsandpatenttrolls.pdf.

Christopher Anthony Cotropia and Mark A Lemley. Copying in patent law. *North Carolina Law Review*, 87:1421, 2009.

Angelique Beek-Van Doremaele and Niels Smetsers. Netherlands. In *Tax Incentives on Research and Development (R&D)* Int [2015 Basel Congress], pages 481–503.

Lisa Evers, Helen Miller, and Christoph Spengel. Intellectual property box regimes: effective tax rates and tax policy considerations. *International Tax and Public Finance*, 22(3):502–530, jun 2014. doi: 10.1007/s10797-014-9328-x.

Robin Feldman and Mark A Lemley. Do patent licensing demands mean innovation? *Iowa Law Review*, 101:137–189, 2015.

Jane G Gravelle. A patent/innovation box as a tax incentive for domestic research and development. Technical Report R44522, Congressional Research Service, June 2016a.

Jane G Gravelle. Policy options to address profit shifting: Carrots or sticks? *Tax Notes*, pages 121–134, July 2016b.

Rachel Griffith, Helen Miller, and Martin O'Connell. Ownership of intellectual property and corporate taxation. *Journal of Public Economics*, 112:12–23, 2014. doi: https://doi.org/10.1016/j.jpubeco.2014.01.009.

Zvi Griliches. Research costs and social returns: Hybrid corn and related innovations. *Journal of Political Economy*, 66(5):419–431, 1958.

Harry Grubert. Intangible income, intercompany transactions, income shifting, and the choice of location. *National Tax Journal*, 56:221–42, March 2003.

Bronwyn H Hall, Jacques Mairesse, and Pierre Mohnen. Measuring the returns to r&d. Working Paper 15622, National Bureau of Economic Research, December 2009.

Tax Incentives on Research and Development (R&D), volume 100A of *Cahiers de Droit Fiscal International*, 2015 Basel Congress. International Fiscal Association.

Adam B Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598, 1993. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2118401.

Todd R Kaplan and David Wettstein. The possibility of mixed-strategy equilibria with constant-returns-to-scale technology under Bertrand competition. *Spanish Economic Review*, 2(1):65–71, 2000. doi: 10.1007/s101080050018.

Ben Klemens. Valuing patents and trademarks in complex production chains. *SSRN Electronic Journal*, December 2015. doi: 10.2139/ssrn.2702102.

Guy Klemens. *The cellphone : the history and technology of the gadget that changed the world*. McFarland, Jefferson, N.C, 2010.

Borbála Kolozs and Annamária Kőszegi. Hungary. In *Tax Incentives on Research and Development (R&D)* Int [2015 Basel Congress], pages 365–382.

Harold William Kuhn and Albert William Tucker, editors. *Contributions to the Theory of Games*. Number 28 in Annals of Mathematical Studies. Princeton University Press, 1953.

Mark A Lemley. The myth of the sole inventor. *Michigan Law Review*, 110(5): 709–760, 2012. URL http://www.jstor.org/stable/23216799.

Richard Levitan and Martin Shubik. Price duopoly and capacity constraints. *International Economic Review*, 13(1):111–122, 1972. URL http://www.jstor.org/stable/2525908.

Stavroula Malla, Richard Gray, and Peter Phillips. Gains to research in the presence of intellectual property rights and research subsidies. *Review of Agricultural Economics*, 26(1):63–81, 2004.

Peter R Merrill, James R Shanahan Jr, José Elías Tomé Gómez, Guillaume Glon, Paul Grocott, Auke Lamers, Diarmuid MacDougall, Alina Macovei, Rémi Montredon, Thierry Vanwelkenhuyzen, Alexandru Cernat, Stephen Merriman, Rachel Moore, Gregg Muresan, Pieter Van Den Berghe, and Andrea Linczer. Is it time for the united states to consider the patent box? *Tax Notes*, pages 1665–1675, March 2012. URL http://www.pwc.com/us/en/washington-national-tax/assets/merrill0326.pdf.

Felipe Alonso Murillo. Spain. In *Tax Incentives on Research and Development (R&D)* Int [2015 Basel Congress], pages 679–702.

William D. Nordhaus. Schumpeterian profits in the american economy: Theory and measurement. Working Paper 10433, National Bureau of Economic Research, April 2004. URL http://www.nber.org/papers/w10433.

Thomas Plümper, Vera E Troeger, and Hannes Winner. Why is there no race to the bottom in capital taxation? *International Studies Quarterly*, 53(3):761–786, 2009. URL http://www.jstor.org/stable/27735120.

Katy Rabindran and Emma Walsh. United kingdom. In *Tax Incentives on Research and Development (R&D)* Int [2015 Basel Congress], pages 789–812.

Paul M Romer. Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102, 1990.

Michael Rustad. *Global Internet Law*. Hornbook. West Academic, 2016.

Gale Shapley. *A value for n-person games*, pages 307–318. In Kuhn and Tucker [1953], 1953.

Stephen E Shay, Jr J Clifton Fleming, and Robert J Peroni. R&D tax incentives: Growth panacea or budget trojan horse? *Tax Law Review*, Forthcoming, 2016.

Robert M Solow. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65, feb 1956. doi: 10.2307/1884513.

Neal Stender, Joel Stark, and Forrest L Ye. China income tax preferences for high/new-tech enterprises (HNTE). Tax law update, Orrick, Herrington & Sutcliffe LLP, January 2010.

Annette Alstadsæter, Salvador Barrios, Gaëtan Nicodème, Agnieszka Maria Skonieczna, and Antonio Vezzani. Patent boxes design, patents location and local r&d. Technical report, 2015.

J G Tewksbury, M S Crandall, and W E Crane. Measuring the societal benefits of innovation. *Science*, 209(4457):658–662, August 1980. doi: 10.1126/science.209.4457.658.

Charles M Tiebout. A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424, 1956.

Antony Ting. iTax: Apple's international tax structure and the double non-taxation issue. *British Tax Review*, (1), March 2014.

Catherine Tucker. Patent trolls and technology diffusion. *SSRN Electronic Journal*, March 2013. doi: 10.2139/ssrn.2136955.

John D Wilson. A theory of interregional tax competition. *Journal of Urban Economics*, 19(3):296–315, May 1986. doi: 10.1016/0094-1190(86)90045-8.

Henry Bosley Woolf, editor. *Webster's Collegiate Dictionary*. G & C Merriam Company.

George R Zodrow and Peter Mieszkowski. Pigou, Tiebout, property taxation, and the underprovision of local public goods. 1986.