Patent exhaustion regime and international production sharing: Winners and losers?

Olena Ivus and Edwin L.-C. Lai*

Abstract

On May 30, 2017, the Supreme Court held that the initial authorized sale of a patented item outside the U.S., just as one within the U.S., "exhausts" all rights of the patentee to that item under the Patent Act. This decision goes against the Government's position that a foreign sale authorized by the U.S. patentee should exhaust U.S. patent rights by default unless the patentee expressly reserved those rights.

In a simple North-South model, we examine how a shift in Northern patent regime from presumptive international exhaustion (PIE)—which presumes that a foreign sale triggers exhaustion but permits express-reservation of rights—to absolute international exhaustion (AIE)—which effectively precludes express-reservation of rights—impacts global welfare and its distribution. PIE subjects the firms that source in the South and sell in the North to the risk of a patent infringement lawsuit but allows the firms to price discriminate internationally. AIE allows the firms to avoid a patent infringement lawsuit but also precludes international price discrimination. Firms differ in the extent to which they engage in international fragmentation of their production process.

We find that being engaged in the international fragmentation of production is neither a necessary nor a sufficient condition for a firm to be more likely to prefer AIE over PIE. Nonetheless, it is still true that Northern firms with high cost share of components that can be outsourced to the South tend to prefer AIE, while firms with low cost share of those components tend to prefer PIE. A shift from PIE to AIE increases fragmentation, lowers production costs, and eliminates North-South price discrimination in the final good. Global welfare rises. Welfare also rises in the North but falls in the South if the impact on firms' incentive for production fragmentation is weak, cost savings from sourcing in the South are low, or the Northern and Southern markets are more dissimilar in size and in their average willingness to pay for most goods.

Keywords: Intellectual property rights; patent exhaustion doctrine; production fragmentation

JEL classification: F10; O34

---

*Olena Ivus (corresponding author): Smith School of Business, Queen's University, 143 Union Street, Kingston, ON K7L 3N6, Canada, oivus@business.queensu.ca. Edwin L.-C. Lai: Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, elai@ust.hk, Edwin.l.lai@gmail.com.
We thank Jay Choi, Taiji Furusawa, Jay Kesan, Zorina Khan, Adam Mossoff, Walter Park, Michael Risch, Mark Schultz, seminar participants at Hitotsubashi University, the 2017 EPIP conference, and the Thomas Edison Innovation Fellowship Meeting in Austin, Texas for helpful comments and discussions. We especially thank Ted Sichelman for his invaluable comments.

1 Introduction

In the realm of intellectual property protection and regulation, the exhaustion doctrine (also referred to as the "first-sale" doctrine) represents a limit on IP-owner rights. In the context of patent law, the doctrine of patent exhaustion implies that the initial authorized sale of a patented item—e.g., by a manufacturer directly authorized by the patent owner—"exhausts" any rights of the patentee to seek payment from downstream buyers. In the United States, effecting the judge-made patent exhaustion doctrine (i.e., determining if the exhaustion doctrine is applicable) has proven controversial. In instances where patent owners sell individual components of a multi-component patented product, the sale of an incomplete article or component which does not completely practice or embody the patent in suit may or may not trigger the exhaustion doctrine. Also, in some instances, upstream patent owners attempt to contractually restrict downstream buyers in order to preserve their right to collect additional license fees. Whether and when such downstream contractual restrictions can overcome the doctrine of patent exhaustion has been the subject of many conflicting, and often vague, judicial decisions over the past century. Further complicating the matters is the distinction between national and international patent exhaustion regime, which is important when a patented item is sold globally. Even if exhaustion applies to authorized sales within a country, it does not necessarily apply to authorized sales abroad.

The agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs) recognizes the exhaustion doctrine, but has left resolution of the matter to individual members. So far in the U.S., the exhaustion doctrine's application and scope has been determined by case law. In *Quanta Computer, Inc. v. LG Electronics, Inc.* (2008) the Supreme Court extended the scope of exhaustion as applied to domestic sales, holding that patent exhaustion is triggered by, among other things, an authorized sale of a component that "substantially embodies" the patent in question or is a "material part" of the patented invention—even if it does not completely practice the patent—such that the only reasonable and intended use of the component is to be finished under the terms of the patent. However, the Supreme Court's decision did little to dispel the confusion about what constitutes "authorization" or whether a patent owner can impose contractual restrictions on downstream buyers to limit patent exhaustion. Regarding the scope of exhaustion as applied to foreign sales, since the 2001 decision in *Jazz Photo Corp. v. International Trade Commission*, the Federal Circuit Court has held that patent exhaustion does not apply internationally, and that importers who attempted to sell foreign-purchased items within the U.S. would be liable for patent infringement. The Supreme Court has not addressed international patent exhaustion in over 125 years—since Boesch v. Gräff (1890) until *Lexmark Int'l, Inc. v. Impression Products, Inc.* (2017)—but has held in *Kirtsaeng v. John Wiley & Sons, Inc.* (2013) that an authorized foreign sale of copyrighted works exhausts U.S. copyright. Although *Kirtsaeng* is the copyright decision, it created uncertainty about the limits of patent exhaustion as it remained unclear how it relates to a patent owner's U.S. rights to control importation of foreign-purchased products patented in the U.S.

Most recently, in *Lexmark Int'l, Inc. v. Impression Products, Inc.* (2017), the Supreme Court reconsidered the scope of patent exhaustion, and the stakes have never been higher. Impression Products purchased printer cartridges from Lexmark's customers both inside and outside the U.S., but always subject to contractual restrictions as to Impression's reuse or resale of those products. In violation of those restrictions, Impression altered and resold the products in the U.S., undercutting Lexmark and its domestic wholesalers in the process. Two specific questions were presented in

Lexmark regarding patent exhaustion: (1) do the contractual restrictions imposed by Lexmark on Impression's purchase override the doctrine of patent exhaustion, so as to make Impression liable for patent infringement?; and (2) even if patent exhaustion applies to purchases made by Impression within the United States, does it apply to Impression's purchases *outside* of the United States, so that Impression can import the cartridges into the United States and resell without being liable for patent infringement?

Lexmark's case against Impression attracted the attention of a number of major players in patent-dependent industries. On one hand were firms like Intel, LG Electronics, Samsung and others reliant on patent-protected inputs to produce their own products. These firms argued that their products include components from numerous suppliers, and that allowing upstream patent owners to limit exhaustion via contractual restrictions would require them to "trace the patent rights of every component it purchases and then negotiate appropriate license arrangements with the component manufacturer (as well as any sub-component manufacturer)." Opponents to a broader scope of patent exhaustion included the Pharmaceutical Research and Manufacturers of America (PhRMA), the Biotechnology Industry Organization (BIO), and the Intellectual Property Owners Association (IPO). These organizations represent firms and individuals that depend substantially on patent rents, and have argued that patent exhaustion should remain limited so as to allow firms to engage in domestic and international price discrimination, particularly without fear of cheaply sold imports emerging in U.S. markets and undercutting their domestic sales.

In a decision released on May 30, 2017, the Supreme Court ruled in Impression's favor, holding that (1) a patentee's decision to sell a product exhausts all of its patent rights in that item, regardless of any restrictions the patentee purports to impose; and (2) an authorized sale outside the United States, just as one within the United States, exhausts all rights under the Patent Act. With this decision, the Supreme Court overturned the Federal Circuit's ruling that upstream patent owners may opt out of patent exhaustion via contractual restrictions, and that patent rights are not exhausted by an authorized sale abroad (even where no reservation of rights accompanies the sale). This decision also goes against the Government's position that a foreign sale authorized by the U.S. patentee should exhaust U.S. patent rights by default unless the patentee expressly reserved those rights.[1]

U.S. patent exhaustion regime is undoubtedly a serious and very contentious policy issue. It carries significant consequences for firms' licensing activity, trade, and welfare. Yet the actual policy implications of patent exhaustion and the *Lexmark* decision are far from clear. This paper assesses the social welfare implications of a regime of international patent exhaustion in the setting of global

[1]The Supreme Court's decision left open the possibility of contractual enforcement. If a licensee has exceeded the limitations in a license agreement directly executed between the patent owner and the licensee, the patent owner may be able to assert a contract claim. Nonetheless, patent owners tend to favor patent infringement claims over contract claims. First, patent claims offer a wider array and scope of remedies than contract claims, particularly injunctive relief. Also, a contract claim usually affords no consequential or punitive damages. Contract cases have the additional disadvantage of being tried in foreign courts, and the cost of contract is higher due to overseas actions. Finally, to enforce downstream restrictions in contract, "privity"—that is, some direct contractual relationship between the parties—is required, and this either precludes suit or substantially raises contracting costs. For instance, in *Lexmark*, the patent owner could not assert a contractual claim against the alleged infringer, because the alleged infringer purchased the products-at-issue from the patent owner's customers rather than patent owner itself.

production sharing. More specifically, it examines how a shift in patent regime from presumptive international exhaustion (PIE)—which presumes that a foreign sale triggers exhaustion but permits express-reservation of rights—to absolute international exhaustion (AIE)—which effectively precludes express-reservation of rights—impacts international transactions and the fragmentation of the production process and with that, global welfare and its distribution.[2]

We develop a simple North-South model in which consumers in each region have vertical preferences and differ in their willingness to pay for each good. Across the two regions, Northern consumers have higher average willingness to pay for each good, and these regional differences in the pattern of demand provide firms with an incentive to price discriminate internationally. The production of a good combines two technologically-sophisticated components, one of which is "global" and can be sourced from the North or the South. The global component contains many different technology-intensive subcomponents, some of which include a patented feature. The patent owner of all patents embodied in this component is located in the North and licenses the manufacturing plants in the North or the South to make the component using its patents with a contractual restriction that the component is only sold to firms for use in the production of a good destined for the domestic market. Whether or not this restriction preserves the patent owner's Northern rights to patents embodied in the component depends on the Northern regime of patent exhaustion.

We assume that absolute patent exhaustion applies within the North, so that the licensing of the manufacturing plants in the North exhausts all patent holder's rights to the component.[3] Exhaustion also applies to the licensing of the plants in the South in a regime of AIE. In a regime of PIE, on the other hand, exhaustion applies to the licensing of the plants in the South unless the patent holder imposes contractual restrictions on downstream buyers via the manufacturer, in which case a firm attempting to use the global component purchased in the South in the production of a good destined for the North must enter into a license with the patent holder or risk patent infringement in the North. The firm that sources the global component in the South knows that its purchase of the component in the South is subject to the contractual restriction but it does not know at the time of production of the final good whether this contractual restriction overrides the doctrine of patent exhaustion so that the firm is liable for patent infringement when selling the good in the North. The firm is uncertain about whether it has the legal right to use the patents embodied in the global component for producing a good to be sold in the North because contractual provisions appear incomplete; and gathering this information and, if necessary, obtaining the legal rights is too costly for the firm, as it requires the firm to trace the patent rights of every subcomponent contained in the technologically sophisticated global component (i.e., determine which subcomponents in the global component are within the scope of a valid and enforceable patent and whether the patentee has reserved the Northern rights) and then negotiate appropriate license fees with the patent holder(s). When international exhaustion is presumptive, the firm that

[2]The Supreme Court's decision left open some potential loopholes, such as converting sales transactions to pure "license/lease" transactions and restricting the "authorization" of the manufacturer to sell the product. However, it is unclear to what extent lower courts will allow patent holders to effectively bypass absolute exhaustion through these potential loopholes, so we generally assume throughout the paper that they do not apply. To the extent they do apply, then such situations are essentially akin to a presumptive exhaustion regime.

[3]Ivus, Lai, and Sichelman (2017) offers the first formal closed-economy model of *national* patent exhaustion that incorporates transaction costs in consumer licensing, and examines how a shift in patent policy from absolute to presumptive national exhaustion affects social welfare.

sources in the South runs the risk of a patent infringement lawsuit in the North. If successfully sued, the court will require the firm to pay damages based on the "reasonable royalty" per unit.

Importantly, we assume a continuum of industries exists, each containing one firm. Firms differ across industries in the cost share of the global component. A Northern firm weighs the savings from sourcing in the South against the risk of patent infringement when selling the final good in the North. This decision process leads to the division of the continuum of industries into those that tend to fragment their production process internationally and those that tend not to. By ranking industries according to their cost share of the global component, we find a threshold industry such that Northern firms in the industries above this threshold tend to source in the South, because the expected cost savings are high, and those below the threshold tend to source from the North, because the risk of a patent lawsuit is too high.

In this context, the firms' preferences over Northern patent exhaustion regime depend on their engagement in international fragmentation of production. PIE subjects the firms that source in the South and sell in the North to the risk of being sued for patent infringement but at the same time, allows the firms to price discriminate internationally. AIE, in contrast, allows the firms to avoid a patent infringement lawsuit but also precludes international price discrimination, as it does not forbid parallel trade.

The theory shows that Northern firms in industries with high cost share of the global component tend to prefer AIE, while firms with low cost share of the global component tend to prefer PIE. Importantly, and somewhat surprisingly, we find that being engaged in the international fragmentation of production is neither a necessary nor a sufficient condition for a firm to be more likely to prefer absolute over presumptive international exhaustion: under certain circumstances, a firm that tends to source in the South under PIE may be more likely to prefer PIE; under other circumstances, a firm that tends not to source from the South under PIE may be more likely to prefer AIE. A shift in Northern patent regime from PIE to AIE would increase fragmentation, lower production costs, and eliminate North-South price discrimination in the final good. Global consumer surplus and welfare would rise. Consumer surplus and welfare would also rise in the North. However, they both fall in the South if the impact on firms' incentive for production fragmentation is weak, cost savings from sourcing in the South are low, or the Northern and Southern markets are more dissimilar in size and in their average willingness to pay for most goods.

Researchers have advanced our understanding of patent exhaustion's welfare implications, though theoretical literature on the subject remains sparse.[4] This paper's contribution to the literature is threefold. First, the paper studies the regime of presumptive international exhaustion, as compared to the regime of absolute international exhaustion. This distinction has been the focus of recent legal decisions and case comments but has not been seen in the economic literature before. The economic literature in turn drew a distinction between the regime of national exhaustion—under which patent rights are exhausted upon the initial authorised domestic sale, but without limiting patent rights in the item sold abroad—and the regime of international exhaustion—under which patent rights are exhausted upon the initial authorised sale anywhere in the world. Under both

[4]Maskus (2000, 2012) provide an in-depth discussion of the international structure of patent exhaustion policies. Saggi (2017) offers a comprehensive survey of the economic literature on this subject.

of these regimes, the exhaustion of patent rights is mandatory; further contractual restrictions upon the sale of protected items cannot be enforced through patent infringement suits. Under the presumptive rule, by contrast, contractual restrictions can negate patent exhaustion. In the face of high transaction costs (including information costs) in inserting the restrictive provision, interpreting the patent scope and litigating, the presumptive rule creates uncertainty over the scope of patent exhaustion and introduces legal uncertainty for patent users.

Second, this paper underscores the implications of exhaustion policies for the legality of cross-border sourcing of intermediate components and the use of these components in the manufacturing of final products for sale around the world. The economic literature has abstracted from these issues and has focused instead on parallel trade in final goods, interpreting the importing nation's treatment of patent exhaustion as determining the legality of parallel imports.[5] Malueg and Schwartz (1994) view parallel imports as a mechanism for defeating international third-degree price discrimination. When parallel trade is permitted, opportunities to price discriminate are eroded by arbitrage. Maskus and Chen (2004) study parallel imports in the context of the vertical relationships in the international marketing of products, showing that parallel imports can arise when intellectual property rights holders engage in vertical price control in unsegmented markets.[6] Whether or not parallel trade in final goods is permitted also affects the firms' pricing decisions in Valletti (2006) and Saggi (2013), with pricing being uniform under international exhaustion and discriminatory under national exhaustion.[7] In our model, pricing also varies across the two regimes—it is uniform under AIE and discriminatory under PIE,—but patent exhaustion regime also affects the firms' incentive for international fragmentation of production. When a patentee can expressly reserve its rights and in so doing, override the doctrine of patent exhaustion, firms sourcing intermediate goods in the South may be liable for patent infringement in the North, and we incorporate the risk of a patent infringement lawsuit into our model.

Third major modeling contribution of our paper is to study the welfare implications of patent exhaustion regime in the setting of international production sharing with a range of industries having different characteristics. This is crucial, since it allows us to explain why firms differ across industries in their regime preferences. It also allows us to incorporate the differential industry responses (regarding pricing and sourcing decisions) into our welfare analysis. Earlier literature has entirely abstracted from industry differences in exhaustion regime preferences. Saggi (2013)

---

[5] Countries following the regime of national exhaustion do not permit parallel imports.

[6] Ganslandt and Maskus (2007), Raff and Schmitt (2007), and Maskus and Stähler (2014) also develop vertical pricing models of parallel imports.

[7] Several papers have explored the relationship between openness to parallel imports and innovation. Valletti (2006) and Valletti and Szymanski (2006) are among the first papers to discuss the welfare implications of parallel imports from both an *ex-ante* and *ex-post* perspective. Valletti (2006), for example, shows that parallel trade reduces investment into product quality *ex-ante* when differential pricing is demand-based, but leads to higher investment *ex-ante* when differential pricing based on differences in the marginal cost of serving various markets. Li and Maskus (2006) and Li and Robles (2007), for example, introduce an R&D stage into the vertical pricing model of Maskus and Chen (2004). Li and Maskus (2006) finds that parallel imports discourage investment in process innovation, but Li and Robles (2007) shows that parallel imports might actually increases the incentive for product innovation. Grossman and Lai (2008) studies the incentives for product innovation in a setting where the South determines its price control policy in response to the North's policy of patent exhaustion and shows that the Northern policy of international exhaustion, which permits parallel imports, can increase the incentives for product innovation in the North it prompts the South to set a more lenient price control.

suggested that exhaustion policies can be fine-tuned at the level of the industry, but did not explore this idea further.

Our setting is similar to Valletti (2006), in that we adopt a vertically differentiated model with demand-based differential pricing.[8] As Valletti (2006), we also assume price discrimination does not open new markets. Valletti (2006) notes that this assumption tends to bias results against price discrimination. The same caveat applies in our model, but with one additional consideration: the risk of patent litigation under PIE offsets the benefit of price discrimination for a firm. Firms value price discrimination to different degrees, and some are willing to forgo the ability to price discriminate in order to avoid a patent infringement lawsuit. In this framework, the static efficiency of a regime of patent exhaustion depends on the extent of production fragmentation and the North-South difference in production costs, market size, and consumer tastes.

In our paper, the firm's incentive to fragment production internationally is a major determinant of the welfare implications of a regime. In this respect, our paper is related to Saggi (2013), where the focus is on the firm's incentive to export. Saggi (2013) shows that the firm's export decision, which is endogenously determined by the Northern policy of patent exhaustion and the Southern policy of IPRs protection, is a major determinant of equilibrium policy choices and their welfare effects.

The paper proceeds as follows. Section 2 develops the model and discusses the Northern regimes of presumptive and absolute international exhaustion. Section 3 establishes the equilibrium sourcing decisions under each regime and describes firms' regime preference. We assess the welfare implications of the Northern patent regimes in Section 4 and summarise our findings in Section 5. Section 6 provides a brief conclusion. All necessary derivations and proofs are contained in the Appendix.

2 Model Set-up

There exists a continuum of industries indexed by $z \in [0, 1]$. There is one firm in each industry which produces one good. The world is comprised of two regions indexed by i: a developed North ($i = N$) and a developing South ($i = S$). For industry z, there are $L_N(z)$ consumers in the North and $L_S(z)$ consumers in the South. Without loss of generality, we assume that Northern market is larger: $L_N(z) > L_S(z)$ for all z, and that the wage rate is w_N in the North and w_S in the South, such that $w_N > w_S$. We normalize the Northern wage rate by setting $w_N = 1$. Thus, $w_S \in (0, 1)$.

[8] In Valletti (2006), price discrimination can arise from either differences in consumer demand or differences in the marginal cost of serving various markets. The regime of international patent exhaustion raises welfare *ex-post* but lowers investment *ex-ante* when differential pricing is demand-based, but has opposite impact (i.e., lowers welfare *ex-post* but raises investment *ex-ante*) when differential pricing is cost-based.

2.1 Preferences

For each industry, consumers in each region have vertical preferences and differ in their type ϕ, which determines their willingness to pay for each good.[9] A consumer with higher ϕ has higher willingness to pay for each good. We assume that ϕ is uniformly distributed over the interval $[0, \Phi_N(z)]$ in the North and the interval $[0, \Phi_S(z)]$ in the South. Therefore, within each industry, the probability density of consumers per type is $1/\Phi_i(z)$, where $i = \{N, S\}$. The Northern consumers have higher willingness to pay for each good on average: $\Phi_N(z) > \Phi_S(z)$ for all z.

The utility that a type-ϕ consumer in region i can obtain from her purchase and consumption of good z is given by:

$$U_i(\phi, z) = h(z)\phi - p_i(z),$$

where $h(z)$ is a positive parameter which signifies the dispersion in the willingness to pay for good z among all consumers of z and $p_i(z)$ is the price of good z in region i.

A consumer of type $\phi > p_i(z)/h(z)$ purchases one unit of good z. A consumer of type $\phi < p_i(z)/h(z)$ does not purchase good z and earns zero utility. Thus, demand for good z in region i is linear and given by the number of consumers in that region who purchase good z:

$$q_i(z) = L_i(z) \int_{\frac{p_i(z)}{h(z)}}^{\Phi_i(z)} \left[\frac{1}{\Phi_i(z)}\right] d\phi = \frac{L_i(z)}{\Phi_i(z)} \left[\Phi_i(z) - \frac{p_i(z)}{h(z)}\right]. \tag{1}$$

2.2 Production

Production of final good z combines two intermediate goods: c1 and c2.[10] c1 and c2 are different for different z. We refer to these intermediate goods as 'components' although depending on the industry, these goods could also be called 'ingredients.' Without loss of generality, we use z to denote the cost share of c1 in the total cost of good z. In other words, the quantity of good z produced is given by $y(z) = A(z)\left[x_1(z)\right]^z \left[x_2(z)\right]^{1-z}$, where $A(z)$ is a constant for any given z such that $A(1) = 1$; $x_1(z)$ is the quantity of c1 and $x_2(z)$ is the quantity of c2 used in the production. Assembly of the final good z from c1 and c2 is costless. Production of c2 is technologically demanding, and the North's productivity advantage in producing c2 is large enough compared with the South to ensure that c2 is produced only in the North. For simplicity of exposition, we assume the firm producing the final good z also produces c2. Production of c1 is less technologically demanding and so, c1 can be produced in both the South and the North by manufacturing firms under perfect competition.

The production of c1 and c2 requires labour, some non-patented subcomponents that can be purchased from the competitive market, and a number of patented subcomponents. Without loss of generality, assume that all patents for the subcomponents contained in c2 are owned by the firm producing the good z. The firm develops and produces the patented subcomponents in c2,

[9] The model is a vertically differentiated model as in Shaked and Sutton (1982).

[10] The result would be qualitatively the same if we assume that there are more than two intermediate goods, the production of some of which can be outsourced to the South.

purchases the non-patented subcomponents in c2 from the competitive market, and then assembles c2 using labour. The unit labour cost for developing and producing the patented subcomponents in c2, producing the non-patented subcomponents in c2 and assembling c2 in the North equals $\beta(z)\, w_N = \beta(z)$. The firm producing the good z purchases c1 from the open competitive market. All patents for the patented subcomponents in c1 are owned by another patent holder located in the North. The patented subcomponents in c1 is a "material part" of this patent holder's invention.[11] The patent holder develops and produces the patented subcomponents in c1 and licenses the manufacturing plants in the North or the South to make c1 using its subcomponents and in return, receives a license fee of $\lambda(z)$ per unit of c1. In addition to the payment of the license fee to the owner of patents in the patented subcomponents of c1, the producer of c1 needs to pay the labor cost in producing the non-patented subcomponents and assembly of c1. The unit labour cost for producing the non-patented subcomponents in c1 and assembling c1 equals $\gamma(z)\, w_N = \gamma(z)$ in the North and $\gamma(z)\, w_S$ in the South. This implies that the total unit cost of producing c1 equals $\lambda(z) + \gamma(z)$ in the North and $\lambda(z) + \gamma(z)\, w_S$ in the South. Intermediate good c2 is technologically more sophisticated than c1 and this is reflected in its higher Northern unit production cost: $\beta(z) > \lambda(z) + \gamma(z)$ for all z.

The unit cost of producing good z equals $c_N(z) = B(z)\, [\lambda(z) + \gamma(z)]^z\, [\beta(z)]^{1-z}$ when c1 is sourced in the North and $c_S(z) = B(z)\, [\lambda(z) + \gamma(z)\, w_S]^z\, [\beta(z)]^{1-z}$ when c1 is sourced in the South, where $B(z) = \left[z^z (1-z)^{1-z}\right]^{-1} / A(z)$.

2.3 Patent exhaustion regime

The Northern firms' pricing and sourcing decisions depend on the Northern regime of patent exhaustion.[12] In line with the Supreme Court's ruling in *Lexmark*, we assume that absolute patent exhaustion applies nationally. That is, an authorized sale within the North of an item that is a "material part" of the patented invention exhausts all patentees's rights to that item, regardless of any restrictions the patentee purports to impose.[13] Regarding international patent exhaustion, two legal regimes are possible in the North:

***Presumptive International Exhaustion (PIE)*: In a Northern regime of PIE, an authorized foreign sale of an item that is a "material part" of the patented invention exhausts all patent holder's rights to that item unless the patent holder imposes contractual restrictions on downstream buyers via the manufacturer. If the patent holder imposes such restrictions, downstream buyers attempting to sell foreign-purchased items in the North must enter into a license with the patent holder or risk patent infringement in the North.

***Absolute International Exhaustion (AIE)*: In a Northern regime of AIE, an authorized

[11]The result would be qualitatively the same if we assume that the patent rights of the subcomponents are owned by many patent holders.

[12]For simplicity, we assume there is no patenting activity in the South and so, the Southern regime of patent exhaustion does not play a role here.

[13]The Supreme Court has held that "If the patentee negotiates a contract restricting the purchaser's right to use or resell the item, it may be able to enforce that restriction as a matter of contract law, but may not do so through a patent infringement lawsuit."

foreign sale of an item that is a "material part" of the patented invention exhausts all patent holder's rights to that item.

We assume that in a regime of PIE, the owner of patents in c1 licenses the manufacturing plants in the South to make c1 using its patents with a contractual restriction that c1 is only sold to firms for use in the production of a good destined for the Southern market. Under this restriction, a firm attempting to use c1 in the production of a good destined for the Northern market must negotiate appropriate royalty rate with the patent holder or risk patent infringement in the North.

The firm that sources c1 in the South knows that its purchase of c1 in the South is subject to the contractual restriction but it does not know at the time of production of good z whether this contractual restriction overrides the doctrine of patent exhaustion so that the firm is liable for patent infringement when selling the good in the North. The firm is uncertain about whether it has the legal right to use the patents embodied in c1 for producing a good to be sold in the North because contractual provisions appear incomplete. Gathering this information and, if necessary, obtaining the legal rights is too costly for the firm, as it requires the firm to trace the patent rights of every subcomponent contained in the technologically sophisticated component c1 (i.e., determine which subcomponents in c1 are within the scope of a valid and enforceable patent and whether the patentee has reserved the Northern rights) and then negotiate appropriate royalty rate(s) with the patent holder(s).

When international exhaustion is presumptive, the firm that sources c1 in the South and sells good z in the North runs the risk of a patent infringement lawsuit in the North, because the sale of c1 to the firm is not expressly authorised by the patent holder given the contractual restriction. But at the same time, PIE allows the firm to price-discriminate internationally. We assume that under PIE, the firm producing the good z,—which owns all patents for the subcomponents in c2,—subjects the sale of the final good to an express no-resale abroad restriction (thus forbidding parallel trade). AIE, in contrast, allows the firm to avoid a patent infringement lawsuit but also prevents the firm from engaging in international price discrimination, as it does not forbid parallel trade and the firm's price discrimination incentive is eroded by a perfect and costless arbitrage.

2.4 Sourcing decision

The firm has to decide which region to source c1 from. This decision is non-trivial when the North adopts PIE. Sourcing c1 in the South permits the cost savings arising from lower Southern wages but under PIE, it also carries the risk of a patent infringement lawsuit when selling the good z in the North. We assume that with probability Ω, the ownwer of all patents in c1 will sue the firm for patent infringement and will win the litigation, i.e., the patents will be held valid and considered infringed by the firm. If successfully sued, the firm will be required by the court to pay damages based on the "reasonable royalty" $r(z)$ per unit. Sourcing c1 from the North eliminates the risk of litigation, but precludes the cost savings from lower wages in the South.

Under AIE, sourcing c1 in the South carries no risk of patent litigation but still permits the cost savings from lower wages in the South.[14]

3 Equilibrium

The model has two stages. In Stage I, the firm decides whether to source c1 from the North or the South. The firm develops and produces c2 in the North, and assembles good z in the North. In Stage II, the firm sets the price(s) for the good z and sells it in both markets. The model is solved using backward induction.

3.1 Equilibrium sourcing decision under PIE

Under PIE, the firm producing good z chooses two prices, $p_N(z)$ and $p_S(z)$. The prices are set for a given unit cost of good z depending on the sourcing decision made in Stage I. Price in region i is set to maximize profit in that region, which is given by

$$\Pi_i(z) = [p_i(z) - c(z)]q_i(z) = [p_i(z) - c(z)]\frac{L_i(z)}{\Phi_i(z)}\left[\Phi_i(z) - \frac{p_i(z)}{h(z)}\right],$$

where $c(z) = c_N(z)$ when c1 is sourced in the North and $c(z) = c_S(z)$ when c1 is sourced in the South, and the second equality follows from (1). Differentiating $\Pi_i(z)$ with respect to $p_i(z)$ and setting the result to zero, we find the profit maximizing price in industry z and region i:

$$p_i(z) = \frac{1}{2}[h(z)\Phi_i(z) + c(z)].$$

At these price, demand and profit in industry z and region i are as follows:

$$q_i(z) = \frac{L_i(z)}{2h(z)\Phi_i(z)}[h(z)\Phi_i(z) - c(z)] \quad \text{and} \quad \Pi_i(z) = \frac{L_i(z)}{4h(z)\Phi_i(z)}[h(z)\Phi_i(z) - c(z)]^2.$$

For simplicity, we assume $L_S(z) = 1$, $\Phi_S(z) = 1$, $L_N(z) = L(z) > 1$, and $\Phi_N(z) = \Phi(z) > 1$ for all z. With this normalization, the North-South gap in prices is given by $p_N(z) - p_S(z) = h(z)[\Phi(z) - 1]/2$. The price gap is high when the maximum willingness-to-pay in the North is relatively high, i.e., $[\Phi(z) - 1]$ is high, and the dispersion in the willingness to pay for each good, $h(z)$, is high. We also assume that $c(z) < h(z)\ \forall z \in [0, 1]$ so that $q_i(z) > 0\ \forall z \in [0, 1]$.

The firm in industry z has to decide whether to source c1 in the North or the South.

[14]Under AIE, the firm attempting to use c1 sourced in the South in the production of a good destined for the North could be subject to a *contractual* claim if it agrees to indemnify the seller in the South for any breach of contract with the patent holder. But under PIE, in addition to the contract claim, the patent owner may be able to assert a patent infringement claim; and as discussed in footnote 1, patent owners tend to favour patent infringement claims. As such, legal uncertainty for patent users is higher under PIE than under AIE. For simplicity, we abstract from contractual actions and focus on patent infringement actions instead.

When c1 is sourced in the North, the firm's global sales and profit in industry z are as follows:

$$q^{dN}(z) = q_N(z) + q_S(z) = \frac{1}{2h(z)} \left\{ h(z)[L(z)+1] - \left[1 + \frac{L(z)}{\Phi(z)}\right] c_N(z) \right\}; \tag{2}$$

$$\Pi^{dN}(z) = \Pi_N(z) + \Pi_S(z) = \frac{1}{4h(z)} \left\{ \frac{L(z)}{\Phi(z)} [h(z)\Phi(z) - c_N(z)]^2 + [h(z) - c_N(z)]^2 \right\}; \tag{3}$$

where the superscript dN denotes "differential" pricing—made possible by the express no-resale abroad restriction—when c1 is sourced in the North.

When c1 is sourced in the South, the firm's global sales are given by:

$$q^{dS}(z) = q_N(z) + q_S(z) = \frac{1}{2h(z)} \left\{ h(z)[L(z)+1] - \left[1 + \frac{L(z)}{\Phi(z)}\right] c_S(z) \right\}, \tag{4}$$

where the superscript dS denotes "differential" pricing when c1 is sourced in the South.

A firm that sources c1 in the South faces the risk of patent litigation when selling the good z in the North. With probability Ω, the owner of all patents in c1 will sue the firm for patent infringement and the court will require the firm to pay damages based on the "reasonable royalty" $r(z)$ per unit. According to Shapiro (2010), a natural benchmark level for reasonable royalty under patent law is the royalty rate or a fixed license fee that would be negotiated between the patent holder and the firm based on the assumption that (i) the firm were aware of patents (for the subcomponents in c1) when making its production decision and (ii) the patents were known to be valid.

To find $r(z)$, we assume symmetric Nash Bargaining between the patent holder and the firm over a fixed license fee $R(z)$ and then divide the fee by the number of units sold $q^{dS}(z)$. If negotiations are successful (i.e., the two parties agree on a license fee), the firm will source c1 in the South and produce the good z at the unit cost of $c_S(z)$. The payoff to the firm from accepting a license at a fixed rate $R(z)$ is given by $\hat{\Pi}^{dS}(z) - R(z)$, where

$$\hat{\Pi}^{dS}(z) = \frac{1}{4h(z)} \left\{ \frac{L(z)}{\Phi(z)} [h(z)\Phi(z) - c_S(z)]^2 + [h(z) - c_S(z)]^2 \right\}. \tag{5}$$

The patent holder's payoff from this license is $R(z)$. If the two parties do not agree on a license fee, the firm will source c1 in the North and its payoff will be equal to $\Pi^{dN}(z)$ in (3), while the patent holder's payoff will be zero. The two parties' combined agreement payoff is thus $\hat{\Pi}^{dS}(z)$ and their combined disagreement payoff is $\Pi^{dN}(z)$. With equal bargaining power, the parties will share equally the gains from the bargain, which implies $R(z) = 0.5[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)]$. The reasonable per unit royalty rate is thus $r(z) = 0.5[\tilde{\Pi}^{dS}(z) - \Pi^{dN}(z)]/q^{dS}(z)$, and the size of the damages awarded to the patent holder for infringement is $r(z)q^{dS}(z)$. Assuming further that in the event of litigation, the firm will incur the litigation cost of C, we find that the firm's profit when sourcing c1 in the South is given by $\Pi^{dS}(z) = \hat{\Pi}^{dS}(z) - \Omega[r(z)q^{dS}(z) + C]$, which simplifies to:

$$\Pi^{dS}(z) = \left(1 - \frac{\Omega}{2}\right) \hat{\Pi}^{dS}(z) + \frac{\Omega}{2} \Pi^{dN}(z) - \Omega C. \tag{6}$$

Let B, λ, γ, β, L, h, and Φ denote random variables that represent the *ex-ante* values of $B(z)$, $\lambda(z)$, $\gamma(z)$, $\beta(z)$, $L(z)$, $h(z)$, and $\Phi(z)$ respectively. In other words, the variables $B(z)$, $\lambda(z)$, $\gamma(z)$, $\beta(z)$, $L(z)$, $h(z)$, and $\Phi(z)$ are the *ex-post* realizations of the respective random variables. *Ex-ante*, we only know that the distribution of each random variable is identical and independent across all z, but we do not know their realizations. Therefore, *ex-ante*, we can express the expected values of $\Pi^{dS}(z)$ and $\Pi^{dN}(z)$ in terms of the above random variables.

Using (6), we find that $E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right] > 0$ iff the following is true:

$$\left(1 - \frac{\Omega}{2}\right) E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] - \Omega C > 0. \tag{7}$$

We refer to Ω as the "patent strength" in the North. The inequality (7) implies that the firm in industry z will choose to source in the South if the patent-strength-adjusted gain in profit from sourcing in the South exceeds the expected litigation cost. Using (2), (3) and (5), we find that:

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] = E\left\{\frac{\widetilde{v}(z)}{4h}\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)\left[\widetilde{c}_N(z) + \widetilde{c}_S(z)\right]\right]\right\}, \tag{8}$$

where $\widetilde{c}_N(z) \equiv B(\lambda+\gamma)^z(\beta)^{1-z}$, $\widetilde{c}_S(z) \equiv B\left(\lambda + \gamma w_S\right)^z \left(\beta\right)^{1-z}$ and $\widetilde{v}(z) \equiv \widetilde{c}_N(z) - \widetilde{c}_S(z)$ are random variables that denote *ex-ante* expressions of $c_N(z)$, $c_S(z)$ and $v(z)$. In other words, the equation (8) states that the *ex-ante* expected profit gain from sourcing in the South equals the *ex-ante* expected value of the unit cost reduction from sourcing c1 in the South, $\widetilde{v}(z)$, times the global sales.

Since $\beta(z) > \lambda(z) + \gamma(z)$ for all z, the *ex-ante* expected unit cost of producing good z falls with an increase in z whether c1 is sourced in the North or the South. Furthermore, the expected value of the unit cost reduction from sourcing c1 in the South, $\widetilde{v}(z)$, increases with z if the following sufficient condition holds: $1 + \ln\lambda(z) > \ln\beta(z)$ for all z. Lemma 1 establishes this result.

Lemma 1 *Assume the parameters $\lambda(z)$, $\gamma(z)$, and $\beta(z)$ satisfy $1 + \ln\lambda(z) > \ln\beta(z) > \ln[\lambda(z) + \gamma(z)]$ for all z. Under this sufficient condition, the expected values of the unit costs of production $\widetilde{c}_N(z)$ and $\widetilde{c}_S(z)$ fall monotonically but the expected value of the unit cost reduction from sourcing c1 from the South $\widetilde{v}(z)$ rises monotonically as z rises, for all z.*

Proof: see Appendix.

From Lemma 1, $dE\left[\widetilde{c}_N(z)\right]/dz < 0$, $dE\left[\widetilde{c}_S(z)\right]/dz < 0$, and $dE\left[\widetilde{v}(z)\right]/dz > 0$. This implies that the *ex-ante* expected profit gain from sourcing in the South rises as z rises. Lemma 2 establishes this result.

Lemma 2 $E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right]$ *rises monotonically with z.*

Proof: see Appendix.

The $z = 0$ firm will source in the North, since $\widetilde{c}_N(0) = \widetilde{c}_S(0)$ and so, $E\left[\Pi^{dS}(z) - \Pi^{dN}(z) \mid z = 0\right] < 0$. The sourcing decision of a $z > 0$ firm depends on the Southern wage w_S, the patent strength Ω, and the litigation cost C.

Define $F(\Omega, w_S) \equiv E\left[\Pi^{dS}(z) - \Pi^{dN}(z) \mid z = 1\right]$. That is, $F(\Omega, w_S)$ is the expected value of $\Pi^{dS}(z) - \Pi^{dN}(z)$, given that $z = 1$. Thus,

$$F(\Omega, w_S) \equiv E\left\{\frac{\widetilde{v}(z)}{4h}\left(1 - \frac{\Omega}{2}\right)\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)[\widetilde{c}_N(z) + \widetilde{c}_S(z)]\right] - \Omega C \mid z = 1\right\}.$$

The $F(\Omega, w_S) = 0$ schedule solves (7) with equality when $z = 1$. Along this schedule, the patent-strength-adjusted expected profit from sourcing in the South is equal to the expected profit from sourcing in the North for the $z = 1$ firm.

Define $\bar{z}$ by $E[\Pi^{dS}(z) - \Pi^{dN}(z) \mid z = \bar{z}] = 0$. Since $E[\Pi^{dS}(z) - \Pi^{dN}(z) \mid z = 0] < 0$ and $dE[\Pi^{dS}(z) - \Pi^{dN}(z)]/dz > 0$ from Lemma 2, we infer that

$$E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right] > 0 \text{ iff } z > \bar{z}.$$

Further define the cutoff $\bar{\Omega}$ implicitly by $F(\bar{\Omega}, 0) = 0$. Proposition 1 follows.

Proposition 1 *Suppose the North adopts PIE. There exists a unique equilibrium with $0 < \bar{z} \leq 1$, where firms with $z \in (0, \bar{z}]$ source c1 in the North and firms with $z \in (\bar{z}, 1]$ source c1 in the South iff the patent strength Ω, the litigation cost C, and the South wage rate w_S are sufficiently low (as shown in the shaded area in Figure 1) so that $F(\Omega, w_S) > 0$.*

Proof: see Appendix.

Proposition 1 states that a unique interior $\bar{z}$ exists if the patent strength, the cost of litigation, and the wage rate in the South are sufficiently low. In this case, firms in the range $(\bar{z}, 1]$ will source c1 in the South, because their cost share of c1 in the total cost of the good z is high enough to make Southern sourcing most profitable. As Ω falls, the patent-strength-adjusted profit from sourcing in the South rises while the expected litigation cost falls and new firms, with lower z, start sourcing in the South.

Figure 1 shows the region of Ω and w_S for which a unique equilibrium with $0 < \bar{z} \leq 1$ exists. The $F(\Omega, w_S) = 0$ schedule intersects the horizontal axes at $w_S = 1$ and the vertical axes at $\Omega = \bar{\Omega}$. The vertical intercept is given by:[15]

$$\bar{\Omega} = \left\{\frac{1}{2} + \frac{C}{E[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z) \mid z = 1]}\right\}^{-1} > 0, \qquad \text{where} \tag{9}$$

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z) \mid z = 1\right] = E\left\{\frac{\gamma}{4h}\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)(2\lambda + \gamma)\right]\right\}. \tag{10}$$

A unique interior $\bar{z}$ exists for any point below $F(\Omega, w_S) = 0$, where $F(\Omega, w_S) > 0$. For example, if the probability of a lawsuit is zero (i.e., $\Omega = 0$), then there exists a unique interior $\bar{z}$ for any w_S. At any point above $F(\Omega, w_S) = 0$, where $F(\Omega, w_S) < 0$, there does not exist an interior $\bar{z}$ and all firms source c1 in the North. The cut-off $\bar{\Omega}$ falls monotonically with C.

[15] It is clear that $\bar{\Omega} > 0$ since $2h(1+L) - (1 + \frac{L}{\Phi})[2\lambda + \gamma] > 0$ by assumption. Furthermore, $\bar{\Omega} > 1$ when the litigation cost is $C < E[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)]/2$, and $\bar{\Omega} < 1$ when the litigation cost is $C > E[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)]/2$.

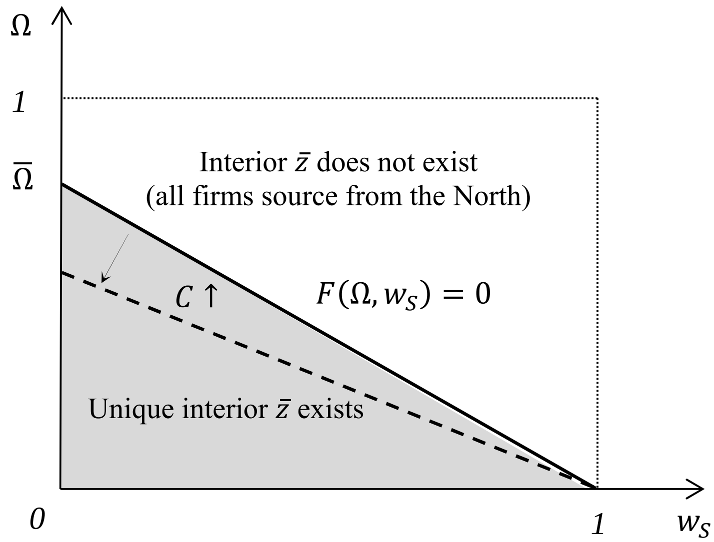

Figure 1: **The existence of a unique equilibrium with $0 < \bar{z} \leq 1$ under PIE**

3.2 Equilibrium sourcing decision under AIE

Under AIE, pricing is uniform: $p_i(z) = p^U(z)$. There is no risk of litigation ($\Omega = 0$) and so, all firms source c1 in the South and $c(z) = c_S(z)$. The global profit in industry z is given by:

$$\Pi^U(z) = [p^U(z) - c_S(z)] \left\{ L(z) + 1 - \left[1 + \frac{L(z)}{\Phi(z)}\right] \frac{p^U(z)}{h(z)} \right\},$$

and so, the profit maximizing uniform price is equal to:

$$p^U(z) = \frac{1}{2} \left\{ [h(z)\, \Phi(z) \left[\frac{L(z)+1}{L(z)+\Phi(z)}\right] + c_S(z) \right\}. \tag{11}$$

At this price, global demand and profit in industry z are as follows:[16]

$$q^U(z) = \frac{1}{2h(z)} \left\{ h(z)[L(z)+1] - \left[1 + \frac{L(z)}{\Phi(z)}\right] c_S(z) \right\}; \tag{12}$$

$$\Pi^U(z) = \frac{1}{4h(z)} \left[\frac{\Phi(z)}{L(z)+\Phi(z)}\right] \left\{ h(z)[L(z)+1] - \left[1 + \frac{L(z)}{\Phi(z)}\right] c_S(z) \right\}^2. \tag{13}$$

In each industry $z \in (\bar{z}, 1]$, where firms always source c1 in the South, the optimal uniform price is a weighted average of the discriminatory prices:

$$p^U(z) = \alpha(z) p_N(z) + [1 - \alpha(z)]\, p_S(z), \quad \text{where} \quad \alpha(z) = \frac{L(z)}{L(z)+\Phi(z)} \quad \text{and} \quad z \in (\bar{z}, 1]. \tag{14}$$

[16]Demand in each region is as follows:

$$q_N^U(z) = \frac{1}{2h(z)} \frac{L(z)}{\Phi(z)} \left\{ h(z)\Phi(z) \left[1 + \frac{\Phi(z)-1}{L(z)+\Phi(z)}\right] - c_S(z) \right\}; \quad q_S^U(z) = \frac{1}{2h(z)} \frac{L(z)}{\Phi(z)} \left\{ h(z) \left[1 - \frac{L(z)[\Phi(z)-1]}{L(z)+\Phi(z)}\right] - c_S(z) \right\}.$$

Together with $p_N(z) - p_S(z) = h(z)[\Phi(z) - 1]/2$, this implies that $p^U(z) - p_N(z) = -[1 - \alpha(z)]h(z)[\Phi(z) - 1]/2$ and $p^U(z) - p_S(z) = \alpha(z)h(z)[\Phi(z) - 1]/2$. As Northern regime shifts from PIE to AIE, price falls in the North and rises in the South. When the Northern market is large (i.e., $L(z)$ is high), the weight on the Northern price, $\alpha(z)$, is high and so, the change in the Northern price is relatively small. But when the maximum willingness-to-pay in the North is relatively high (i.e., $\Phi(z)$ is high), the weight on the Northern price is low and so, the change in the Northern price is relatively large.

In each industry $z \in (0, \bar{z}]$, where firms source c1 in the North under PIE and in the South under AIE, the optimal uniform price is a weighted average of the discriminatory prices net of the difference in the production cost of good z, given by $v(z) \equiv c_N(z) - c_S(z)$:

$$p^U(z) = \alpha(z)p_N(z) + [1 - \alpha(z)]\, p_S(z) - \frac{1}{2}v(z), \quad \text{where} \quad z \in (0, \bar{z}].$$

This implies that $p^U(z) - p_N(z) = -\{[1 - \alpha(z)]h(z)[\Phi(z) - 1] + v(z)\}/2$ and $p^U(z) - p_S(z) = \{\alpha(z)h(z)[\Phi(z) - 1] - v(z)\}/2$. As Northern regime shifts to AIE, the production cost of good z falls across firms with $z \in (0, \bar{z}]$. The cost reduction contributes to a price decline in the North and counteracts a price increase in the South.

Furthermore, from (2), (4), and (12), $q^U(z) = q^{dS}(z)$ and $q^U(z) > q^{dN}(z)$ for any $z > 0$. Thus as Northern regime shifts to PIE, global sales rise across firms with $z \in (0, \bar{z}]$ because their production cost falls. For firms with $z \in (\bar{z}, 1]$, the cost of production and global sales remain unchanged.[17]

3.3 Firms' regime preference

A firm $z \in (0, \bar{z}]$, which sources in the North under PIE, will prefer AIE if $E[\Pi^U(z) - \Pi^{dN}(z)] > 0$, i.e., if its expected global profit from sourcing in the South with uniform pricing exceeds its expected global profit from sourcing in the North with differential pricing. From (3) and (13), this is equivalent to

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > E\left\{\frac{h}{4}\alpha\,[\Phi - 1]^2\right\}, \tag{15}$$

where $E[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)]$ is in (8) and $\alpha = L/(L + \Phi)$ is the *ex-ante* expression of $\alpha(z)$. On the left hand side (LHS) of (15) is the expected profit gain from AIE due to the reduction in the production cost of good z, which arises as a firm with $z \in (0, \bar{z}]$ switches from sourcing in the North under PIE to sourcing in the South under AIE, when the risk of patent litigation falls to zero. On the right hand side (RHS) of (15) is the expected profit loss from AIE due to the lost ability to price discriminate geographically, which equals $E[\hat{\Pi}^{dS}(z) - \Pi^U(z)]$.

A firm $z \in (\bar{z}, 1]$, which sources in the South under both regimes, will prefer AIE if $E[\Pi^U(z) - \Pi^{dS}(z)] > 0$, i.e., if its expected global profit with uniform pricing exceeds its expected global profit

[17] The finding that global sales are unchanged under the two regimes provided the cost of production is unchanged is in line with the previous research (e.g., Valletti (2006) and Saggi (2013)).

with differential pricing. From (5), (6) and (13), this is equivalent to

$$\frac{\Omega}{2}E\left[\hat{\Pi}^{dS}(z)-\Pi^{dN}(z)\right]+\Omega C>E\left\{\frac{h}{4}\alpha\left[\Phi-1\right]^{2}\right\}, \quad (16)$$

where $E[\hat{\Pi}^{dS}(z)-\Pi^{dN}(z)]$ is in (8). On the LHS of (16) is the expected profit gain from AIE due to the avoidance of the risk of litigation, and on the RHS is the expected profit loss from AIE due to the lost ability to price discriminate geographically, just as in (15).

Lemma 3 *$E\left[\Pi^{U}(z)-\Pi^{dN}(z)\right]$ and $E\left[\Pi^{U}(z)-\Pi^{dS}(z)\right]$ rise monotonically with z, and $dE\left[\Pi^{U}(z)-\Pi^{dN}(z)\right]/dz>dE\left[\Pi^{U}(z)-\Pi^{dS}(z)\right]/dz$ for any given z.*

Proof: see Appendix.

As established earlier, the $z=0$ firm will source in the North. The $z=0$ firm will also prefer PIE, since $\widetilde{c}_N(0)=\widetilde{c}_S(0)$ and so, (15) does not hold. The regime preference of a $z>0$ firm depends on the Southern wage w_S and if the firm is already sourcing in the South, the patent strength Ω and the litigation cost C.

Define $H(\Omega,w_S)\equiv E\left[\Pi^{U}(z)-\Pi^{dS}(z)\ \mid z=1\right]$. Thus,

$$H(\Omega,w_S)\equiv E\left\{\frac{\widetilde{v}(z)}{4h}\frac{\Omega}{2}\left[2h(1+L)-\left(1+\frac{L}{\Phi}\right)\left[\widetilde{c}_N(z)+\widetilde{c}_S(z)\right]\right]+\Omega C-\frac{h}{4}\alpha(\Phi-1)^{2}\ \mid z=1\right\}$$

The $H(\Omega,w_S)=0$ schedule solves (16) with equality when $z=1$, and it represents all combinations of Ω and w_S such that the expected profit gain under AIE is equal to the expected profit loss from AIE for the $z=1$ firm. Define the cutoff $\bar{\bar{\Omega}}$ implicitly by $H(\bar{\bar{\Omega}},0)=0$.

Further define $G(w_S)\equiv E\left[\Pi^{U}(z)-\Pi^{dN}(z)\ \mid z=1\right]$. Using (3) and (13), we obtain:

$$G(w_S)\equiv E\left\{\widetilde{v}(z)\left[2h(1+L)-\left(1+\frac{L}{\Phi}\right)\left[\widetilde{c}_N(z)+\widetilde{c}_S(z)\right]\right]-\alpha h^{2}(\Phi-1)^{2}\mid z=1\right\}.$$

The $G(w_S)=0$ schedule solves (15) with equality when $z=1$. Along this schedule, the expected profit gain from AIE for the $z=1$ firm due to the reduction in the cost of c1 (as the firm switches from sourcing in the North to sourcing in the South) is equal to the expected profit loss from AIE due to the lost ability to price-discriminate geographically.

Proposition 2 follows.

Proposition 2 *If the patent strength Ω is sufficiently high and the South wage rate w_S is sufficiently low, then there exists a unique equilibrium with $0<\bar{\bar{z}}\leq 1$, where firms with $z\in(0,\bar{\bar{z}}]$ tend to prefer PIE and firms with $z\in(\bar{\bar{z}},1]$ tend to prefer AIE. The necessary and sufficient condition for this to be true is: $\lambda(z)$ is sufficiently low, $\gamma(z)$ is less than but sufficiently close to $h(z)$ (as shown in the shaded area in Figure A2) and, for all z,*

$$2[1+L(z)]-\left[1+\frac{L(z)}{\Phi(z)}\right]>\alpha(z)[\Phi(z)-1]^{2}. \quad (17)$$

Proof: see Appendix.

Proposition 2 states that a sufficiently high patent strength Ω and sufficiently low South wage rate w_S are required in order for some firms to prefer AIE over PIE. When Ω is high, the risk of patent litigation is high under PIE, while it is zero under AIE. Low w_S magnifies the gain from eliminating the risk of litigation, as it implies high damages for patent infringement in the event of litigation.

Proposition 2 also requires a sufficiently low $\lambda(z)$, sufficiently high $\gamma(z)$ (more precisely, $\gamma(z)$ less than but sufficiently close to $h(z)$), and the condition (17). This combination is shown in the shaded area in Figure A2 in the appendix. First, when $\lambda(z)$ is low and $\gamma(z)$ is high, firms prefer AIE for a larger range of Ω, because the gain from avoiding the risk of litigation is high. So, the cut-off $\bar{\bar{\Omega}}$ is low. Second, the condition (17) is required. The RHS in (17) is low when the profit loss from AIE due to the inability to price discriminate geographically is low. The geographical price discrimination is not very valuable to the firms when the Northern and Southern markets are similar in size (i.e., $L(z)$ is low) and consumer tastes (i.e., $\Phi(z)$ is low). The LHS in (17) is high when the profit gain from AIE due to the reduction in the cost of production of good z for firms with $0 < z \leq \bar{z}$ or the elimination of the risk of patent litigation and costly patent damages (for firms with $\bar{z} < z \leq 1$) is high. This is the case when the average global sales, i.e., $[q^{dN}(z) + q^{dS}(z)]/2$, are high.

One intuition of Proposition 2 is that, regardless of whether a high-z firm sources in the South under PIE, it gains a lot when the Northern regime shifts to AIE (as it spends a lot on c1) when w_S is sufficiently low and Ω is sufficiently high (thus PIE is unfavorable) and so prefers AIE to PIE despite the loss of the ability to price discriminate geographically.

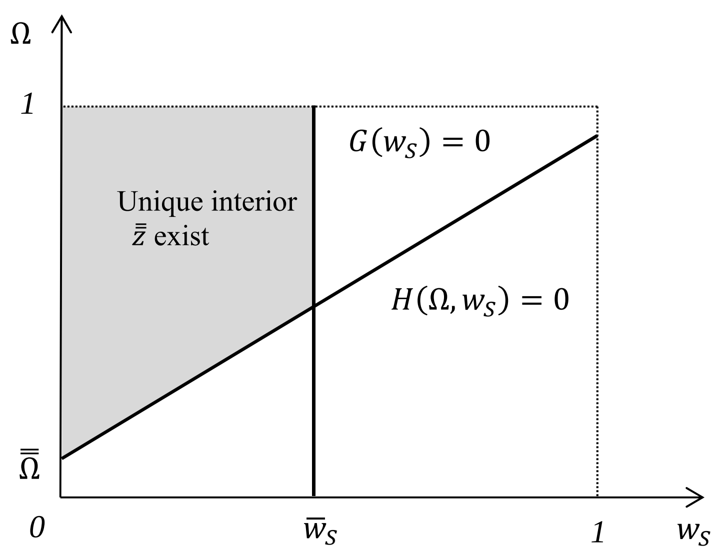

Figure 2: **The existence of a unique equilibrium with** $0 < \bar{\bar{z}} \leq 1$ **under PIE**

Figure 2 shows the region of Ω and w_S for which a unique interior equilibrium with $0 < \bar{\bar{z}} \leq 1$

exists. The $H(\Omega, w_S) = 0$ schedule intersects the vertical axes at $\bar{\bar{\Omega}}$, given by:[18]

$$\bar{\bar{\Omega}} = E\left\{\frac{h}{4}\alpha(\Phi-1)^2\right\}\left\{C + \frac{E[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z) \mid z=1]}{2}\right\}^{-1} > 0, \tag{18}$$

where $E[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z) \mid z = 1]$ is given in (10). The cut-off $\bar{\bar{\Omega}}$ rises as the cost of litigation C falls, but the ratio $\bar{\Omega}/\bar{\bar{\Omega}}$ is independent of C.[19] The $G(w_S) = 0$ schedule intersects the horizontal axes at $w_S = \bar{w}_S$, which is in the range of $(0, 1)$ provided the condition (17) holds.

3.4 Equilibria with two cut-offs

Figure 3 shows the region of Ω and w_S below the $F(\Omega, w_S) = 0$ schedule for which a unique interior equilibrium with $0 < \bar{z} \leq 1$ exists, as explained in Proposition 1. Under the conditions stated in Proposition 2, this region consists of two distinct sub-regions: above and below the $H(\Omega, w_S) = 0$ schedule. For any Ω and w_S in the region below $F(\Omega, w_S) = 0$ and above $H(\Omega, w_S) = 0$, a unique interior equilibrium with $0 < \bar{z} \leq 1$ and $0 < \bar{\bar{z}} \leq 1$ exists.

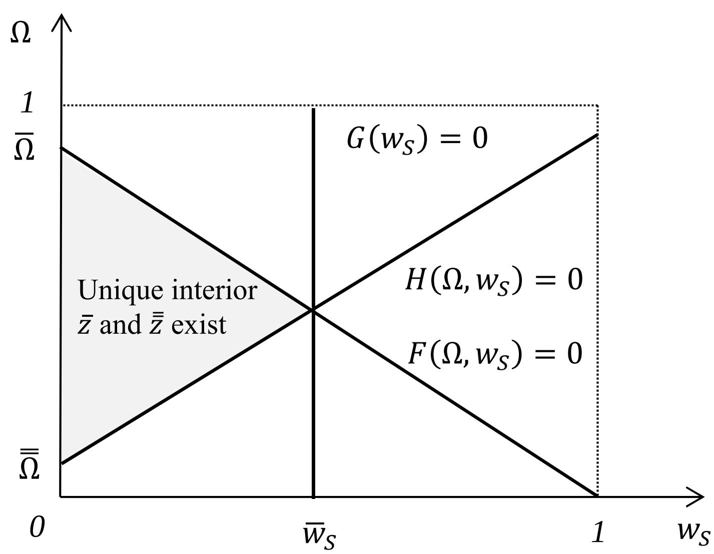

Figure 3: **The existence of a unique equilibrium with** $0 < \bar{z} \leq 1$ **and** $0 < \bar{\bar{z}} \leq 1$

Suppose both interior $\bar{z}$ and $\bar{\bar{z}}$ exist (i.e., $F(\Omega, w_S) > 0$ and $H(\Omega, w_S) > 0$). From Section 3.1, $E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right] > 0$ iff the inequality (7) is true, which is equivalent to:

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > \frac{\Omega C}{1 - \Omega/2},$$

[18]It is clear that $\bar{\bar{\Omega}} > 0$ since $E\left\{2h(1+L) - (1+\frac{L}{\Phi})[2\lambda + \gamma]\right\} > 0$ by assumption.

[19]This ratio is given by $\bar{\Omega}/\bar{\bar{\Omega}} = 4E[\tilde{\Pi}^{dS}(z) - \Pi^{dN}(z) \mid z = 1]/E\{\alpha h(\Phi-1)^2\}$.

where $E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right]$ is given by (8). Further from Section 3.3, $E[\Pi^{U}(z) - \Pi^{dS}(z)] > 0$ iff the inequality (16) holds, which is equivalent to:

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > \frac{E\left\{\frac{h}{4}\alpha\left[\Phi - 1\right]^2\right\} - \Omega C}{\Omega/2}.$$

From Section 3.3, we also know that $E[\Pi^{U}(z) - \Pi^{dN}(z)] > 0$ iff (15) holds.

From the above, we can conclude the following *ex-ante*:

When Ω is sufficiently large, $\bar{\bar{z}} < \bar{z}$. In that case, an industry with $z > \bar{z}$ is expected to source from the South under PIE and prefer AIE if given the choice between PIE and AIE. But there also exist some industry $z \in [\bar{\bar{z}}, \bar{z}]$ that is expected not to source from the South under PIE but prefer AIE. When Ω is sufficiently small, $\bar{z} < \bar{\bar{z}}$. In that case, an industry with $z > \bar{\bar{z}}$ is expected to source from the South under PIE and prefer AIE if given the choice between PIE and AIE. But there also exist some industry $z \in [\bar{z}, \bar{\bar{z}}]$ that is expected to source from the South under PIE but prefer PIE.

Thus, we have Proposition 3.

Proposition 3 *Suppose there exist unique interior $\bar{z}$ and $\bar{\bar{z}}$ as explained in Propositions 1 and 2 (as shown in the shaded area in Figure 3). Then there exists a unique critical value of the patent strength $\widetilde{\Omega}$ such that $\bar{\bar{\Omega}} < \widetilde{\Omega} < \bar{\Omega}$ and the following is true: (i) $\bar{\bar{z}} = \bar{z}$ iff $\Omega = \widetilde{\Omega}$; (ii) $\bar{\bar{z}} < \bar{z}$ iff $\Omega > \widetilde{\Omega}$; and (iii) $\bar{z} < \bar{\bar{z}}$ iff $\Omega < \widetilde{\Omega}$.*

Proof: see Appendix.

The implication of Proposition 3 is that, in general, the fact that a firm sources c1 in the South under PIE is neither necessary nor sufficient for it to be more likely to favor AIE over PIE. Note that w_S must be sufficiently small ($w_S < \bar{w}_S$) for interior $\bar{z}$ and $\bar{\bar{z}}$ both to exist. If Ω is sufficiently large ($\Omega > \widetilde{\Omega}$), a firm that sources c1 in the South under PIE tends to favors AIE, as its gain from the removal of the risk of patent litigation under AIE is high enough that it trumps its loss from the inability to price discriminate geographically. In fact, some high-z firms that do not source c1 from South under PIE also tend to favor AIE. These firms tend not to source c1 in the South under PIE as the expected patent damages are too high, but since w_S is sufficiently small, they expect to gain a lot from sourcing c1 in South when the North switches to AIE, as they still spend a sufficiently large percentage of their unit cost on c1. This expected cost saving trumps the expected loss due to uniform pricing. Conversely, if Ω is sufficiently small ($\Omega < \widetilde{\Omega}$), a firm that tends to source c1 from the North tends to favor PIE. But some firms with sufficiently low z that tend to source c1 in South under PIE also tend to favor PIE over AIE, as the expected loss due to the inability to price discriminate geographically dominates the expected gain from the removal of the risk of patent litigation when sourcing in the South. Therefore, the fact that a firm tends to source c1 in the South is neither a necessary nor sufficient condition for it to be more likely to prefer AIE over PIE.

Figure 4 shows two equilibria, $0 < \bar{\bar{z}} < \bar{z} < 1$ on the left and $0 < \bar{z} < \bar{\bar{z}} < 1$ on the right, which are possible under the conditions stated in Proposition 2. In the equilibrium with $\bar{\bar{z}} < \bar{z}$, a shift from PIE to AIE lowers the expected profit from $E[\Pi^{dN}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(0, \bar{\bar{z}}]$, and raises the expected profit from $E[\Pi^{dN}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(\bar{\bar{z}}, \bar{z}]$ and from $E[\Pi^{dS}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(\bar{z}, 1]$. In the equilibrium with $\bar{z} < \bar{\bar{z}}$, a shift to AIE lowers the expected profit from $E[\Pi^{dN}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(0, \bar{z}]$ and from $E[\Pi^{dS}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(\bar{z}, \bar{\bar{z}}]$, and raises the expected profit from $E[\Pi^{dS}(z)]$ to $E[\Pi^{U}(z)]$ for the firms in $(\bar{\bar{z}}, 1]$.

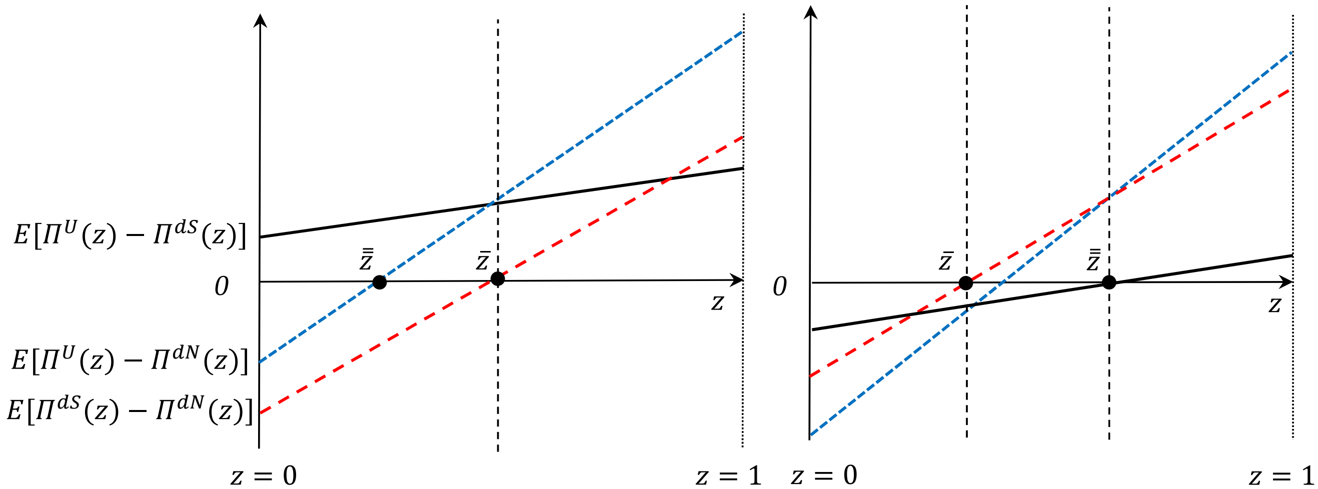

Figure 4: **Equilibria with two cut-offs**

4 Welfare comparison

In this section, we examine the welfare implications of a Northern regime of patent exhaustion. We first analyze how consumer surplus (CS) compares across PIE and AIE and then proceed with a comparison of global welfare and its distribution across the regions. Hereinafter, all references to consumer surplus, producer surplus and welfare are in expected value terms.

4.1 Consumer surplus

Under PIE, global CS in industry z is given by:

$$CS^d(z) = E\left\{ \frac{L}{\Phi} \int_{\frac{\widetilde{p}_N(z)}{h}}^{\Phi} \left[h\phi - \widetilde{p}_N(z) \right] d\phi + \int_{\frac{\widetilde{p}_S(z)}{h}}^{1} \left[h\phi - \widetilde{p}_S(z) \right] d\phi \right\},$$

where a tilde above a variable denotes its *ex-ante* value. The first term is expected CS in the North and the second term is expected CS in the South. Assuming there exists a unique interior $\bar{z}$ as explained in Proposition 1, the differential prices $\widetilde{p}_N(z) = [h\Phi + \widetilde{c}(z)]/2$ and $\widetilde{p}_S(z) = [h + \widetilde{c}(z)]/2$

depend on the production cost as follows: $\widetilde{c}(z) = \widetilde{c}_N(z)$ when $z \in [0, \bar{z}]$ and $\widetilde{c}(z) = \widetilde{c}_S(z)$ when $z \in (\bar{z}, 1]$. Consequently, global CS varies across industries as follows:

$$CS^d(z) = CS^{dN}(z) = E\left\{\frac{1}{8h}\left[\frac{L}{\Phi}[h\Phi - \widetilde{c}_N(z)]^2 + [h - \widetilde{c}_N(z)]^2\right]\right\} \quad \text{for} \quad z \in [0, \bar{z}]; \tag{19}$$

$$CS^d(z) = CS^{dS}(z) = E\left\{\frac{1}{8h}\left[\frac{L}{\Phi}[h\Phi - \widetilde{c}_S(z)]^2 + [h - \widetilde{c}_S(z)]^2\right]\right\} \quad \text{for} \quad z \in (\bar{z}, 1]. \tag{20}$$

Across all industries, global CS under PIE is $CS^d = \int_0^{\bar{z}} CS^{dN}(z)dz + \int_{\bar{z}}^1 CS^{dS}(z)dz$.

Under AIE, global CS is the same in each industry and is given by:

$$CS^U(z) = E\left\{\frac{L}{\Phi}\int_{\frac{\widetilde{p}^U(z)}{h}}^{\Phi}\left[h\phi - \widetilde{p}^U(z)\right]d\phi + \int_{\frac{\widetilde{p}^U(z)}{h}}^{1}\left[h\phi - \widetilde{p}^U(z)\right]d\phi\right\}.$$

Using the *ex-ante* value of $p^U(z)$ in (11), we rewrite $CS^U(z)$ as follows:

$$CS^U(z) = E\left\{\frac{1}{8h}\left(1 + \frac{L}{\Phi}\right)\left[h\Phi\frac{L+1}{L+\Phi} - \widetilde{c}_S(z)\right]^2 + \frac{h}{2}\alpha(\Phi - 1)^2\right\}. \tag{21}$$

Across all industries, global CS under AIE is $CS^U(z) = \int_0^1 CS^U(z)dz$.

We now consider each region individually. As the patent regime shifts from PIE to AIE, CS in region i changes from $CS_i^{dN}(z)$ to $CS_i^U(z)$ in industry $z \in [0, \bar{z}]$ and from $CS_i^{dS}(z)$ to $CS_i^U(z)$ in industry $z \in (\bar{z}, 1]$. The relevant expressions for the North ($i = N$) are as follows:

$$CS_N^{dN}(z) = E\left\{\frac{1}{8h}\frac{L}{\Phi}[h\Phi - \widetilde{c}_N(z)]^2\right\}; \qquad CS_N^{dS}(z) = E\left\{\frac{1}{8h}\frac{L}{\Phi}[h\Phi - \widetilde{c}_S(z)]^2\right\}; \tag{22}$$

$$CS_N^U(z) = E\left\{\frac{1}{8h}\frac{L}{\Phi}\left[h\Phi\frac{L+1}{L+\Phi} - \widetilde{c}_S(z)\right]^2 + \frac{\alpha}{2}(\Phi - 1)[h\Phi - \widetilde{c}_S(z)]\right\}; \tag{23}$$

and for the South ($i = S$), they are given by:

$$CS_S^{dN}(z) = E\left\{\frac{1}{8h}[h - \widetilde{c}_N(z)]^2\right\}; \qquad CS_S^{dS}(z) = E\left\{\frac{1}{8h}[h - \widetilde{c}_S(z)]^2\right\}; \tag{24}$$

$$CS_S^U(z) = E\left\{\frac{1}{8h}\left[h\Phi\frac{L+1}{L+\Phi} - \widetilde{c}_S(z)\right]^2 - \frac{\alpha}{2}(\Phi - 1)[h - \widetilde{c}_S(z)]\right\}. \tag{25}$$

Proposition 4 establishes the results.

Proposition 4 *A shift in patent regime from PIE to AIE increases global consumer surplus: $CS^U > CS^d$. Consumer surplus always rises in the North, $CS_N^U > CS_N^d$. However, consumer* *surplus falls in the South, $CS_S^U > CS_S^d$, under this sufficient condition: $\gamma(z) < h(z)\,\alpha(z)[\Phi(z)-1]$ for all z.*

As the regime shifts from PIE to AIE, global CS rises in each industry z. This is apparent from

$$CS^U(z)-CS^{dN}(z) = E\left\{\frac{3h}{8}\alpha[\Phi-1]^2 + \frac{\widetilde{v}(z)}{8h}\left[2h(1+L)-\left(1+\frac{L}{\Phi}\right)[\widetilde{c}_N(z)+\widetilde{c}_S(z)]\right]\right\} > 0; \quad (26)$$

$$CS^U(z) - CS^{dS}(z) = E\left\{\frac{3h}{8}\alpha[\Phi-1]^2\right\} > 0. \quad (27)$$

Global CS rises across all industries due to two forces at play. First, AIE brings about uniform pricing. The resulting gain in CS is given by the first term on the RHS of (26) and the term on the RHS of (27). Second, AIE increases production fragmentation. The cost of production and price fall in industries $z \in [0, \bar{z}]$ as a result and CS rises. This gain is represented by the second term on the RHS of (26).

Consumer surplus also necessarily rises in each industry in the North:

$$CS_N^U(z) - CS_N^{dN}(z) = E\left\{\frac{\alpha}{8}(\Phi-1)\left[2[h\Phi-\widetilde{c}_S(z)]+h(1-\alpha)(\Phi-1)\right]\right\} +$$

$$+ E\left\{\frac{\widetilde{v}(z)}{8h}\left[2hL-\frac{L}{\Phi}[\widetilde{c}_N(z)+\widetilde{c}_S(z)]\right]\right\} > 0; \quad (28)$$

$$CS_N^U(z) - CS_N^{dS}(z) = E\left\{\frac{\alpha}{8}(\Phi-1)\left[2[h\Phi-\widetilde{c}_S(z)]+h(1-\alpha)(\Phi-1)\right]\right\} > 0. \quad (29)$$

Intuitively, this is because the optimal price necessarily falls in the North. First, the Northern price falls as firms switch to the uniform pricing. The resulting gain in CS in given by the first term on the RHS of (28) and the term on the RHS of (29).[20] Second, in industries $z \in [0, \bar{z}]$, the Northern price also falls as firms start sourcing from the South, thus cutting on their production costs and lowering the Northern price. This gain is given by the second term on the RHS of (28).

Further as we show in the Appendix, CS falls in the South if $\gamma(z) < h(z)\,\alpha(z)[\Phi(z)-1]$ for all z. This follows since the following is true under this sufficient condition:

$$CS_S^U(z)-CS_S^{dN}(z) = E\left\{-\frac{\alpha}{8}(\Phi-1)\left[2[h-\widetilde{c}_S(z)]-h\alpha(\Phi-1)\right] + \frac{\widetilde{v}(z)}{8h}\left[2h-[\widetilde{c}_N(z)+\widetilde{c}_S(z)]\right]\right\} < 0, \quad (30)$$

and it is also true that

$$CS_S^U(z) - CS_S^{dS}(z) = E\left\{-\frac{\alpha}{8}(\Phi-1)\left[2[h-\widetilde{c}_S(z)]-h\alpha(\Phi-1)\right]\right\} < 0. \quad (31)$$

The ambiguity in the impact on the Southern CS stems from the ambiguity in the impact on the Southern optimal price. On one hand, the price rises in the South as firms in each industry forgo

[20] This term equals the expected value of the change in price $\widetilde{p}_N(z) - \widetilde{p}^U(z) > 0$ times the average Northern sales across the two regimes $[\widetilde{q}_N(z) + \widetilde{q}_N^U(z)]/2$ when $\widetilde{c}(z) = \widetilde{c}_S(z)$.

discriminatory pricing. This reduces Southern CS, as represented by the first term on the RHS of (30) and the term on the RHS of (31).[21] But at the same time, the Southern price falls in industries $z \in [0, \bar{z}]$ as firms in these industries reduce their cost of production. The resulting gain in the Southern CS is given by the second term on the RHS of (30). When $\gamma(z) < h(z)\alpha(z)[\Phi(z) - 1]$, the cost of production in $z \in [0, \bar{z}]$ declines only little relative to the increase in the Southern price following the switch to uniform pricing. The Southern CS falls in each industry z as a result.[22] This occurs when the unit labour requirement for producing component c1 is low (i.e., $\gamma(z)$ is low), the dispersion in the willingness to pay for each good is high (i.e., $h(z)$ is high), and the Northern and Southern markets are more dissimilar in size and the average willingness to pay for each good (i.e., $L(z)$ and $\Phi(z)$ are high).

4.2 Global welfare and its distribution

Welfare consists of consumer surplus and producer surplus (PS), where the PS is the sum of the Northern firms' expected profits and the patent holder's expected overall rent from manufacturer licensing as well as patent infringement awards in the event of a successful patent lawsuit.

Under PIE, global welfare in industry z is given by $W^{dN}(z) = CS^{dN}(z) + E\left\{\Pi^{dN}(z) + \lambda(z)\,\widetilde{q}^{dN}(z)\right\}$ when $z \in [0, \bar{z}]$ and $W^{dS}(z) = CS^{dS}(z) + E\left\{\Pi^{dS}(z) + [\lambda(z) + \Omega r(z)]\widetilde{q}^{dS}(z)\right\}$ when $z \in (\bar{z}, 1]$. These expressions simplify to:[23]

$$W^{dN}(z) = E\left\{\frac{3}{8h}\left[\frac{L}{\Phi}[h\Phi - \widetilde{c}_N(z)]^2 + [h - \widetilde{c}_N(z)]^2\right] + \lambda\widetilde{q}^{dN}(z)\right\}; \tag{32}$$

$$W^{dS}(z) = E\left\{\frac{3}{8h}\left[\frac{L}{\Phi}[h\Phi - \widetilde{c}_S(z)]^2 + [h - \widetilde{c}_S(z)]^2\right] - \Omega C + \lambda\widetilde{q}^{dS}(z)\right\}; \tag{33}$$

where $\widetilde{q}^{dN}(z)$ and $\widetilde{q}^{dS}(z)$ are the *ex-ante* values of $q^{dN}(z)$ and $q^{dS}(z)$ in (2) and (4).

Across all industries, global welfare under PIE is $W^d = \int_0^{\bar{z}} W^{dN}(z)dz + \int_{\bar{z}}^1 W^{dS}(z)dz$.

Under AIE, global welfare in each industry is given by:

$$W^U(z) = E\left\{\frac{3}{8h}\left(1 + \frac{L}{\Phi}\right)\left[h\Phi\frac{L+1}{L+\Phi} - \widetilde{c}_S(z)\right]^2 + \frac{h}{2}\alpha(\Phi - 1)^2 + \lambda\widetilde{q}^U(z)\right\}, \tag{34}$$

where $\widetilde{q}^U(z)$ is the *ex-ante* value of $q^U(z)$ in (12).

Across all industries, global welfare under AIE is $W^U = \int_0^1 W^U(z)dz$.

[21] This term equals the expected value of the change in price $\widetilde{p}_S(z) - \widetilde{p}^U(z) < 0$ times the average Southern sales across the two regimes $[\widetilde{q}_S(z) + \widetilde{q}^U_S(z)]/2$ when $c(z) = c_S(z)$.

[22] When $\gamma(z) < h(z)\alpha(z)[\Phi - 1]$, a shift from PIE to AIE leads to a decline in sales in the South in $z \in [0, \bar{z}]$.

[23] We substitute for $\widetilde{q}^{dN}(z)$ and $\widetilde{q}^{dS}(z)$ using the *ex-ante* values of $q^{dN}(z)$ and $q^{dS}(z)$ in (2) and (4); for $\Pi^{dN}(z)$ and $\Pi^{dS}(z)$ using (3) and (6); for $CS^{dN}(z)$ and $CS^{dS}(z)$ using (19) and (20); and use the result that $r(z)\widetilde{q}^{dS}(z) = \hat{\Pi}^{dS}(z) - \Pi^{dN}(z)$, where $\tilde{\Pi}^{dS}(z)$ and $\Pi^{dN}(z)$ are given by (5) and (3).

In the South, welfare equals CS. In the North, welfare consists of CS and PS and is given by $W_N^{dN}(z)$ in $z \in [0, \bar{z}]$ and $W_N^{dS}(z)$ in $z \in (\bar{z}, 1]$ under PIE and $W_N^U(z)$ in $z \in [0,1]$ under AIE, where

$$W_N^{dN}(z) = E\left\{\frac{3}{8h}\frac{L}{\Phi}[h\Phi - \widetilde{c}_N(z)]^2 + \frac{1}{4h}[h - \widetilde{c}_N(z)]^2 + \lambda\widetilde{q}^{dN}(z)\right\}; \tag{35}$$

$$W_N^{dS}(z) = E\left\{\frac{3}{8h}\frac{L}{\Phi}[h\Phi - \widetilde{c}_S(z)]^2 + \frac{2}{8h}[h - \widetilde{c}_S(z)]^2 - \Omega C + \lambda\widetilde{q}^{dS}(z)\right\}; \tag{36}$$

$$W_N^U(z) = E\left\{\left(\frac{3}{8h}\frac{L}{\Phi} + \frac{1}{4h}\right)\left[h\Phi\frac{L+1}{L+\Phi} - \widetilde{c}_S(z)\right]^2 + \frac{\alpha}{2}(\Phi - 1)[h\Phi - \widetilde{c}_S(z)] + \lambda\widetilde{q}^U(z)\right\}. \tag{37}$$

Proposition 5 establishes the results.

Proposition 5 *A shift in patent regime from PIE to AIE increases expected global welfare:* $W^U > W^d$. *Across the two regions, welfare always rises in the North,* $W_N^U > W_N^d$. *However, welfare falls in the South,* $W_S^U > W_S^d$, *under this sufficient condition:* $\gamma(z) < h(z)\alpha(z)[\Phi(z) - 1]$ *for all* z.

A shift in patent regime from PIE to AIE increases global welfare in each industry z:

$$W^U(z) - W^{dN}(z) = E\left\{\frac{h}{8}\alpha(\Phi - 1)^2 + \frac{3\widetilde{v}(z)}{8h}\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)[\widetilde{c}_N(z) + \widetilde{c}_S(z)]\right] + \right.$$

$$\left. + \lambda\frac{\widetilde{v}(z)}{2h}\left(1 + \frac{L}{\Phi}\right)\right\} > 0; \tag{38}$$

$$W^U(z) - W^{dS}(z) = E\left\{\frac{h}{8}\alpha(\Phi - 1)^2 + \Omega C\right\} > 0. \tag{39}$$

Here, three effects are at play, all of which contribute to the global welfare gain. First, firms switch from discriminatory to uniform pricing under AIE. This change in the pricing strategy hurts firms, but it benefits consumers even more on average. The global welfare rises in each industry as a result, as represented by the first term in (38) and (39). Second, firms in industries $z \in [0, \bar{z}]$ switch from sourcing in the North to sourcing in the South under AIE. This leads to the reduction in the cost of production and the associated welfare gain is given by the second term in (38). This also leads to the increase in the patent holder's overall rent from manufacturer licensing, as global sales in industries $z \in [0, \bar{z}]$ rise. The associated welfare gain given by the third term in (38). Last, firms in industries $z \in (\bar{z}, 1]$ avoid the risk of a patent infringement lawsuit and the cost of litigation when sourcing in the South under AIE. The last term in (39) shows the associated welfare gain.

Welfare also rises in the North, despite Northern firms' profits fall in industries $z \in [0, \bar{\bar{z}}]$. This result follows from

$$W_N^U(z) - W_N^{dN}(z) = E\left\{\frac{\alpha}{8}(\Phi - 1)\left[h(1-\alpha)(\Phi - 1) + 2[h - \widetilde{c}_S(z)]\right] + \right.$$

$$+ \frac{\widetilde{v}(z)}{4h}\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)[\widetilde{c}_N(z) + \widetilde{c}_S(z)]\right] + \frac{\widetilde{v}(z)}{8h}\left[2hL - \frac{L}{\Phi)}[\widetilde{c}_N(z) + \widetilde{c}_S(z)]\right] +$$

$$+\lambda\frac{\widetilde{v}(z)}{2h}\left(1+\frac{L}{\Phi}\right)\Bigg\} > 0. \tag{40}$$

$$W_N^U(z)-W_N^{dS}(z)=E\left\{\frac{\alpha}{8}(\Phi-1)\Big[h(1-\alpha)(\Phi-1)+2[h-\widetilde{c}_S(z)]\Big]+\Omega C\right\}>0; \tag{41}$$

The first term in (40) and (41) is the gain in the Northern welfare due to uniform pricing, which equals the gain in Northern CS net of the profit loss. The next two terms in (40) represent the Northern welfare gain due to the reduction in the cost of production in industries $z \in [0, \bar{z}]$; and the last term in (40) is the Northern welfare gain due to the increase in the patent holder's overall rent from manufacturer licensing in industries $z \in [0, \bar{z}]$. The last term in (40) is the Northern welfare gain due to the elimination of the risk of patent litigation in industries $z \in (\bar{z}, 1]$.

In the South, welfare falls when $\gamma(z) < h(z)\,\alpha(z)[\Phi(z)-1]$ for all z, since CS falls in the South in this case.

5 Summary of findings

Below we summarize our model's three key predictions:

- ***Prediction 1: Sourcing decision under PIE***

  If the patent strength, the litigation cost, and the South wage rate are sufficiently low, high-z industries which rely more on c1 for the production of their final good will be more likely to source c1 from South and low-z industries which rely less on c1 will be more likely to source c1 from the North.

- ***Prediction 2: Firms' regime preference***

  Suppose that Prediction 1 is true. Then if the South wage rate is sufficiently low and the patent strength is sufficiently high, it is also true that industries with sufficiently high z are more likely to prefer AIE and industries with sufficiently low z are more likely to prefer PIE.

  Furthermore, if the patent strength is above a certain threshold, AIE is expected to be preferred by all firms that source c1 in the South under PIE and also by some firms that source c1 in North under PIE. However, if the patent strength is below that threshold, all firms that source c1 from the North under PIE are expected to prefer PIE, but even some firms that source c1 from South under PIE are expected to prefer PIE. Therefore, the fact that a firm sources c1 from South is neither a necessary nor sufficient condition for it to be more likely to prefer AIE to PIE.

- ***Prediction 3: Welfare comparison***

  A shift in patent regime from PIE to AIE increases global consumer surplus and global welfare. Consumer surplus and welfare also rise in the North. However, they both fall in the South if the impact on firms' incentive for production fragmentation is weak, cost savings from sourcing in the South are low, or the Northern and Southern markets are more dissimilar in size and average willingness to pay for each good.

6 Conclusion

Although there is still a technological gap between the North and the South in terms of the capability to develop new technology, it is increasingly the case that the South is able to produce some intermediate goods or components using patented subcomponents imported from the North. These intermediate goods or components are of sufficiently high quality that they can compete with those produced in the North. As the labor cost is lower in the South, components produced by the South are substantially cheaper. This provides an incentive for firms serving the Northern market to offshore the production of some of the intermediate goods to the South, or equivalently, to source these intermediate goods in the South. However, the Northern regime of international patent exhaustion affects such incentive for fragmentation as the firm that sources a patented component in the South and sells the final good in the North might be subject to the risk of a patent infringement lawsuit in the North when the North adopts PIE. Clearly, different industries have different willingness to source in the South under PIE. We find that when Southern wage is lower, the probability of being sued for patent infringement is lower, or the litigation cost is lower, industries that rely more on the components that can be offshored would in fact be more likely to source those components in the South under a Northern regime of PIE.

What if the North changes its patent regime from presumptive to absolute international exhaustion? Interestingly, we find that the industries that source in the South under PIE may not all tend to favor AIE. In general, the fact that a firm sources in the South under PIE is neither necessary nor sufficient for it to have a higher probability of favoring AIE over PIE. The firm's regime preference depends on the probability of being sued for patent infringement in the North when selling the good produced using South-sourced components under PIE. Nonetheless, it is still true that industries that rely more on components that can be offshored would more likely prefer AIE. This creates tension between two types of industries, one that would lobby for PIE, and another that would lobby for AIE. The recent court cases concerning patent exhaustion in the U.S. would have tremendous implications for these industries. It would also have significant consequence on consumer welfare, as a change of the US regime from PIE to AIE would increase fragmentation, lower production costs, as well as eliminate North-South price discrimination in the final good. These factors would increase global consumer surplus and welfare. Consumer surplus and welfare would also rise in the North. However, they both fall in the South if the impact on firms' incentive for production fragmentation is weak, cost savings from sourcing in the South are low, or the Northern and Southern markets are more dissimilar in size and in the average willingness to pay for most goods.

References

[1] Ganslandt, M. and Maskus, K.E., 2007. Vertical distribution, parallel trade, and price divergence in integrated markets. *European Economic Review*, 51(4): 943-970.

[2] Grossman, G.M. and Lai, E.L.-C., 2008. Parallel imports and price controls. *The Rand Journal of Economics*, 39(2): 378-402.

[3] Ivus, O., Lai, E.L.-C. and Sichelman, T.M., 2017. An Economic Model of Patent Exhaustion. *San Diego Legal Studies Paper No.* 17-265. Available at SSRN: `https://ssrn.com/abstract=2921443`

[4] Li, C. and Maskus, K.E., 2006. The impact of parallel imports on investments in cost-reducing research and development. *Journal of International Economics*, 68(2): 443-455.

[5] Li, C. and Robles, J., 2007. Product innovation and parallel trade. *International Journal of Industrial Organization*, 25(2): 417-429.

[6] Malueg, D.A. and Schwartz, M., 1994. Parallel imports, demand dispersion, and international price discrimination. *Journal of International Economics*, 37(3-4): 167-195.

[7] Maskus, K.E., 2012. *Private rights and public problems: the global economics of intellectual property in the 21st century*. Peterson Institute.

[8] Maskus, K.E., 2000. Parallel imports. *The World Economy*, 23(9): 1269-1284.

[9] Maskus, K.E. and Chen, Y., 2004. Vertical price control and parallel imports: theory and evidence. *Review of International Economics*, 12(4): 551-570.

[10] Maskus, K.E. and Stähler, F., 2014. Retailers as agents and the limits of parallel trade. *European Economic Review*, 70: 186-196.

[11] Raff, H. and Schmitt, N., 2007. Why parallel trade may raise producers' profits. *Journal of International Economics*, 71(2): 434-447.

[12] Saggi, K., 2017. Trade, intellectual property rights, and the world trade organization. In: Bagwell, K. and R.W. Staiger (Ed.), Handbook of Commercial Policy. North Holland, 1 edition, pp. 433-512.

[13] Saggi, K., 2013. Market power in the global economy: the exhaustion and protection of intellectual property. *The Economic Journal*, 123(567): 131-161.

[14] Shaked, A. and Sutton, J., 1982. Relaxing price competition through product differentiation. *The Review of Economic Studies*, 49(1): 3-13.

[15] Shapiro, C., 2010. Injunctions, Hold-Up, and Patent Royalties. *American Law and Economics Review*, 12: 280-318.

[16] Valletti, T.M., 2006. Differential pricing, parallel trade, and the incentive to invest. *Journal of International Economics*, 70(1): 314-324.

[17] Valletti, T.M. and Szymanski, S., 2006. Parallel trade, international exhaustion and intellectual property rights: a welfare analysis. *The Journal of Industrial Economics*, 54(4): 499-526.

Appendix

PROOF OF LEMMA 1

Since $\widetilde{c}_N(z) = B[\lambda+\gamma]^z[\beta]^{1-z}$ and $\widetilde{c}_S(z) = B[\lambda+\gamma w_S]^z[\beta]^{1-z}$ and $\beta(z) > \lambda(z)+\gamma(z)$ for all z, we have:

$$dE\left[\widetilde{c}_N(z)\right]/dz = E\left[d\widetilde{c}_N(z)/dz\right] = E\left\{B\left[\ln(\lambda+\gamma)-\ln\beta\right]\widetilde{c}_N(z)\right\} < 0; \tag{A1}$$

$$dE\left[\widetilde{c}_S(z)\right]/dz = E\left[d\widetilde{c}_S(z)/dz\right] = E\left\{B\left[\ln(\lambda+\gamma w_S)-\ln\beta\right]\widetilde{c}_S(z)\right\} < 0. \tag{A2}$$

We next show that $E\left[\widetilde{v}(z)\right] \equiv E\left[\widetilde{c}_N(z)-\widetilde{c}_S(z)\right]$ increases with z for all $z\in[0,1]$. Since $\widetilde{c}_N(0) = \widetilde{c}_S(0)$, and $\{\ln[\lambda(z)+\gamma(z)]-\ln\beta(z)\}$ for all z, is less negative than $\{\ln[\lambda(z)+\gamma(z)\,w_S]-\ln\beta(z)\}$ for all z, we have: $dE\left[\widetilde{v}(z)\right]/dz > 0$ at $z=0$. At $z=1$, $dE\left[\widetilde{v}(z)\right]/dz > 0$ iff the following is true for all z:

$$[\lambda(z)+\gamma(z)w_S]\{\ln\beta-\ln[\lambda(z)+\gamma(z)w_S]\} > [\lambda(z)+\gamma(z)]\{\ln\beta(z)-\ln[\lambda(z)+\gamma(z)]\}, \tag{A3}$$

where the terms on each side are positive. Note that the term on the left hand side of (A3) is concave in w_S and is equal to the term on the right hand side when $w_S=1$. Thus (A3) holds if $[\lambda(z)+\gamma(z)w_S]\{\ln\beta(z)-\ln[\lambda(z)+\gamma(z)w_S]\}$ falls as w_S rises from $w_S=0$ for all z, which in turn requires the following condition to hold: $\ln\beta(z) < 1+\ln\lambda(z)$ for all z. Now,

$$\frac{d^2E\left[\widetilde{v}(z)\right]}{dz^2} = E\left[\frac{d^2\widetilde{v}(z)}{dz^2}\right] = E\left\{\left[\ln(\lambda+\gamma)-\ln\beta\right]\frac{d\widetilde{c}_N(z)}{dz}\right\} - E\left\{\left[\ln(\lambda+\gamma w_S)-\ln\beta\right]\frac{d\widetilde{c}_S(z)}{dz}\right\},$$

which is less than zero when $z=0$ and $z=1$, since $\{\ln[\lambda(z)+\gamma(z)]-\ln\beta(z)\}$ is less negative than $\{\ln[\lambda(z)+\gamma(z)\,w_S]-\ln\beta(z)\}$ and $dE\left[\widetilde{c}_N(z)\right]/dz$ is less negative than $dE\left[\widetilde{c}_S(z)\right]/dz$ at $z=0$ and $z=1$. Does there exist any point at $z\in(0,1)$ where $d^2E\left[\widetilde{v}(z)\right]/dz^2 = 0$? At this point it must be true that $dE\left[\widetilde{c}_N(z)\right]/dz$ is more negative than $dE\left[\widetilde{c}_S(z)\right]/dz$. That is $dE\left[\widetilde{v}(z)\right]/dz < 0$. But this is violated at point B in Figure A1. Therefore, there does not exist any point at $z\in(0,1)$ where $d^2E\left[\widetilde{v}(z)\right]/dz^2 = 0$, which means that $dE\left[\widetilde{v}(z)\right]/dz$ must be monotonically decreasing in $z\in(0,1)$. This further implies that $dE\left[\widetilde{v}(z)\right]/dz > 0$ for all $z\in(0,1)$ if $\ln\beta(z) < 1+\ln\lambda(z)$ for all z (which implies that $E(\beta) < 1+E(\ln\lambda)$).

PROOF OF LEMMA 2

The result that $dE[\Pi^{dS}(z)-\Pi^{dN}(z)]/dz > 0$ follows from Lemma 1 and

$$E\left[\Pi^{dS}(z)-\Pi^{dN}(z)\right] = \left[1-\frac{\Omega}{2}\right]E\left\{\frac{\widetilde{v}(z)}{4h}\left[2h(1+L)-\left(1+\frac{L}{\Phi}\right)\left[\,\widetilde{c}_N(z)+\widetilde{c}_S(z)\right]\right]\right\}-\Omega C.$$

PROOF OF PROPOSITION 1

The proposition is true iff $E[\Pi^{dS}(0)-\Pi^{dN}(0)] < 0$ and $E[\Pi^{dS}(1)-\Pi^{dN}(1)] \geq 0$. First, it is obvious that $E[\Pi^{dS}(0)-\Pi^{dN}(0)] < 0$, since $\widetilde{v}(0)=0$ and $\Omega>0$. Second, $E[\Pi^{dS}(1)-\Pi^{dN}(1)] \geq 0$ iff $F(\Omega,w_S)\geq 0$, where

$$F(\Omega,w_S) \equiv E\left\{\frac{\widetilde{v}(1)}{4h}\left(1-\frac{\Omega}{2}\right)\left[2h(1+L)-\left(1+\frac{L}{\Phi}\right)\left[\widetilde{c}_N(1)+\widetilde{c}_S(1)\right]\right]-\Omega C\right\}.$$

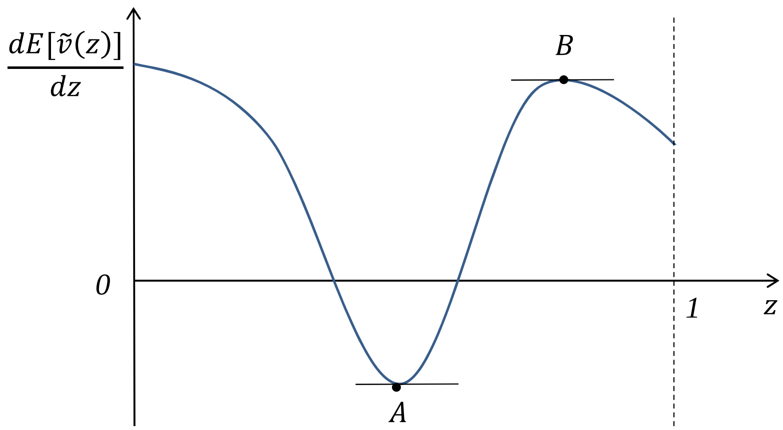

Figure A1: **A violation of $dE[\widetilde{v}(z)]/dz < 0$ at point B**

and $\widetilde{v}(1) \equiv \widetilde{c}_N(1) - \widetilde{c}_S(1)$, $\widetilde{c}_N(1) \equiv [\lambda + \gamma]$ and $\widetilde{c_S}(1) \equiv [\lambda + \gamma w_S]$. In Figure 1, we show that the schedule $F(\Omega, w_S) = 0$ intersects the horizontal axes at $w_S = 1$ and the vertical axes at $\bar{\Omega} > 0$, which is given by:

$$\bar{\Omega} = \left\{ \frac{1}{2} + \frac{C}{E[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)]} \right\}^{-1} > 0, \qquad \text{where} \tag{A4}$$

$$E\left[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)\right] = E\left\{ \frac{\gamma}{4h} \left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)(2\lambda + \gamma) \right] \right\}. \tag{A5}$$

Along the $F(\Omega, w_S) = 0$ schedule, $d\Omega/dw_S < 0$. This result follows since $dF/dw_S < 0$ and $dF/d\Omega < 0$ and so by the implicit function theorem, $d\Omega/dw_S = -[dF/dw_S]/[dF/d\Omega] < 0$. Moreover, since $dF/dw_S < 0$ and $dF/d\Omega < 0$, $F(\Omega, w_S) > 0$ for any combination of Ω and w_S below $F(\Omega, w_S) = 0$. Thus there exists a unique interior $\bar{z}$ for any point below $F(\Omega, w_S) = 0$. For any point above $F(\Omega, w_S) = 0$, there does not exist any interior $\bar{z}$ and all firms source c1 in the North.

PROOF OF LEMMA 3

The results that $dE[\Pi^U(z) - \Pi^{dN}(z)]/dz > 0$ and $dE[\Pi^U(z) - \Pi^{dS}(z)]/dz > 0$ follow from Lemma 1 and

$$E\left[\Pi^U(z) - \Pi^{dN}(z)\right] = E\left\{ \frac{\widetilde{v}(z)}{4h} \left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)[\, \widetilde{c}_N(z) + \widetilde{c}_S(z)] \right] - \frac{h}{4}\alpha[\Phi - 1]^2 \right\};$$

$$E\left[\Pi^U(z) - \Pi^{dS}(z)\right] = E\left\{ \frac{\Omega}{2}\frac{\widetilde{v}(z)}{4h} \left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)[\, \widetilde{c}_N(z) + \widetilde{c}_S(z)] \right] - \Omega C - \frac{h}{4}\alpha[\Phi - 1]^2 \right\}.$$

PROOF OF PROPOSITION 2

Since $\bar{\bar{z}}$ is determined by the intersection of the lower envelope of $E[\Pi^U(z) - \Pi^{dN}(z)]$ and $E[\Pi^U(z) - \Pi^{dS}(z)]$ with the horizontal axis, we need both $E[\Pi^U(1) - \Pi^{dS}(1)] \geq 0$ and $E[\Pi^U(1) - \Pi^{dN}(1)] \geq 0$.

First, $E[\Pi^U(1) - \Pi^{dS}(1)] \geq 0$ iff $H(\Omega, w_S) \geq 0$, where

$$H(\Omega, w_S) \equiv E\left\{\frac{\widetilde{v}(1)}{4h}\frac{\Omega}{2}\left[2h(1+L) - \left(1+\frac{L}{\Phi}\right)[\widetilde{c}_N(1) + \widetilde{c}_S(1)]\right] + \Omega C - \frac{h}{4}\alpha(\Phi-1)^2\right\}$$

and $\widetilde{v}(1) \equiv \widetilde{c}_N(1) - \widetilde{c}_S(1)$, $\widetilde{c}_N(1) \equiv [\lambda + \gamma]$ and $\widetilde{c_S}(1) \equiv [\lambda + \gamma w_S]$.

In Figure 2, we show that the schedule $H(\Omega, w_S) = 0$ intersects the vertical axes at $\bar{\bar{\Omega}} > 0$, which is given by:

$$\bar{\bar{\Omega}} = E\left\{\frac{h}{4}\alpha(\Phi-1)^2\right\}\left[C + \frac{E[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)]}{2}\right]^{-1} > 0, \tag{A6}$$

where $E[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)]$ is given by (A5).

Along the $H(\Omega, w_S) = 0$ schedule, $d\Omega/dw_S > 0$. This result follows since $dH/dw_S < 0$ and $dH/d\Omega > 0$ and so by the implicit function theorem, $d\Omega/dw_S = -[dH/dw_S]/[dH/d\Omega] > 0$. Moreover, since $dH/dw_S < 0$ and $dH/d\Omega > 0$, $H(\Omega, w_S) > 0$ for any combination of Ω and w_S above $H(\Omega, w_S) = 0$. This means a unique interior equilibrium with $0 < \bar{\bar{z}} \leq 1$ exists for any point above $H(\Omega, w_S) = 0$. For any point below $H(\Omega, w_S) = 0$, there does not exist any interior $\bar{\bar{z}}$ and all firms prefer PIE.

Second, $E[\Pi^U(1) - \Pi^{dN}(1)] \geq 0$ is equivalent to $G(w_S) \geq 0$, where

$$G(w_S) \equiv E\left\{\widetilde{v}(1)\left[2h(1+L) - \left(1+\frac{L}{\Phi}\right)[\widetilde{c}_N(1) + \widetilde{c}_S(1)]\right] - \alpha h^2(\Phi-1)^2\right\}.$$

Differentiating $G(w_S)$ with respect to w_S, we find that $dG/dw_S < 0$, since $h(z)[1 + L(z)] - \left[1 + \frac{L(z)}{\Phi(z)}\right][\lambda(z) + \gamma(z)w_S] > 0$ for all z. At $w_S = 1$, we have $G(1) < 0$. Furthermore, as we show below, $G(0) > 0$ at $w_S = 0$ if $\lambda(z)$ is sufficiently low, $\gamma(z)$ is less than but sufficiently close to $h(z)$ and for all z,

$$2[1 + L(z)] - \left[1 + \frac{L(z)}{\Phi(z)}\right] > \alpha(z)[\Phi(z) - 1]^2. \tag{A7}$$

Therefore, there exists a unique interior $0 < \bar{w}_S < 1$, implicitly defined by $G(\bar{w}_S) = 0$, such that $G(w_S) \geq 0$ for any $0 \leq w_S \leq \bar{w}_S$ and $G(w_S) < 0$ for any $\bar{w}_S < w_S \leq 1$. This is shown in Figure 2.

It remains to show that $G(0) > 0$ at $w_S = 0$ if $\lambda(z)$ is sufficiently low, $\gamma(z)$ is less than but sufficiently close to $h(z)$ and for all z, the inequality (A7) holds. We find that $G(0) > 0$ at $w_S = 0$ if the following inequality holds:

$$E\left\{\gamma^2 - 2\gamma\left[h\Phi\frac{L+1}{\Phi+L} - \lambda\right] + h^2[\Phi-1]^2\frac{L\Phi}{[\Phi+L]^2}\right\} < 0. \tag{A8}$$

This inequality holds for a certain set of combinations of *ex-post* realizations of $\lambda(z)$, $\gamma(z)$, $h(z)$, $\Phi(z)$ and $L(z)$, for all z. To find this set, let

$$b \equiv 2\left[h(z)\Phi(z)\frac{L(z)+1}{\Phi(z)+L(z)} - \lambda(z)\right] \quad \text{and} \quad c \equiv h(z)^2[\Phi(z)-1]^2\frac{L(z)\Phi(z)}{[\Phi(z)+L(z)]^2}.$$

Then (A8) holds if $\gamma(z)^2 - b\gamma(z) + c < 0$ holds for all z.

The inequality $\gamma(z)^2 - b\gamma(z) + c < 0$ holds iff $b^2 - 4c > 0$ and $(b - \sqrt{b^2 - 4c})/2 < \gamma(z) < (b + \sqrt{b^2 - 4c})/2$. The parameters $\gamma(z)$ and $\lambda(z)$ must also satisfy $h(z) > \lambda(z) + \gamma(z)$ for the demand in the South to be positive for all industries (in particular for industry $z = 1$ under PIE).

Define $\lambda^*(z)$ as the value of $\lambda(z)$ such that $b^2 - 4c = 0$. Therefore, $\lambda(z) < \lambda^*(z)$ iff $b^2 - 4c > 0$. This defines one allowable region in the $(\lambda(z), \gamma(z))$ space. Next, define

$$\Gamma_1(\lambda(z)) \equiv (b - \sqrt{b^2 - 4c})/2 \quad \text{and} \quad \Gamma_2(\lambda(z)) \equiv (b + \sqrt{b^2 - 4c})/2.$$

Thus, we can define two more allowable regions in the $(\lambda(z), \gamma(z))$ space: $\gamma(z) > \Gamma_1(\lambda(z))$ and $\gamma(z) < \Gamma_2(\lambda(z))$.

It can be easily shown that the curve $\gamma(z) = \Gamma_2(\lambda(z))$ is downward sloping in the $(\lambda(z), \gamma(z))$ space. This is because

$$\frac{\partial \Gamma_2(\lambda(z))}{\partial \lambda(z)} = \frac{1}{2}\frac{\partial b}{\partial \lambda(z)}\left[1 + \frac{b}{\sqrt{b^2 - 4c}}\right] < -2 < 0 \quad \text{as} \quad \frac{\partial b}{\partial \lambda(z)} = -2 \quad \text{and} \quad \frac{b}{\sqrt{b^2 - 4c}} > 1.$$

Moreover, if we define $\gamma_2^*(z) \equiv \Gamma_2(0)$, it can be easily shown that $\gamma_2^*(z) > h(z)$.

Furthermore, it is easy to show that $h(z) - \lambda(z) < (b + \sqrt{b^2 - 4c})/2$ for any $\lambda(z) < \lambda^*(z)$ and so, $\gamma(z) < h(z) - \lambda(z)$ is sufficient for $\gamma(z) < \Gamma_2(\lambda(z))$.

It can also be shown that the curve $\gamma(z) = \Gamma_1(\lambda(z))$ is upward sloping in $(\lambda(z), \gamma(z))$ space. This is because

$$\frac{\partial \Gamma_1(\lambda(z))}{\partial \lambda(z)} = \frac{1}{2}\frac{\partial b}{\partial \lambda(z)}\left[1 - \frac{b}{\sqrt{b^2 - 4c}}\right] > 0.$$

Define $\gamma_1^*(z) \equiv \Gamma_1(0)$. Thus, if we can find a condition under which $\gamma_1^*(z) < h(z)$, then the set of combinations of $\lambda(z)$ and $\gamma(z)$ that satisfy (A8) can be non-empty. This non-empty set is shown in Figure A2. This has to be true for all z.

A necessary and sufficient condition for $\gamma_1^*(z) < h(z)$ is $(b + \sqrt{b^2 - 4c})/2 < h(z)$ at $\lambda(z) = 0$, which is equivalent to:

$$2[1 + L(z)] - \left[1 + \frac{L(z)}{\Phi(z)}\right] > \alpha(z)[\Phi(z) - 1]^2. \tag{A9}$$

Therefore, a sufficient condition for Proposition 2 to hold is that (A9) and that $(\lambda(z), \gamma(z))$ falls in the shaded triangle in Figure A2 for all z.

PROOF OF PROPOSITION 3

Recall the following:

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] = E\left\{\frac{\widetilde{v}(z)}{4h}\left[2h(1+L) - \left(1 + \frac{L}{\Phi}\right)\left[\widetilde{c}_N(z) + \widetilde{c}_S(z)\right]\right]\right\};$$

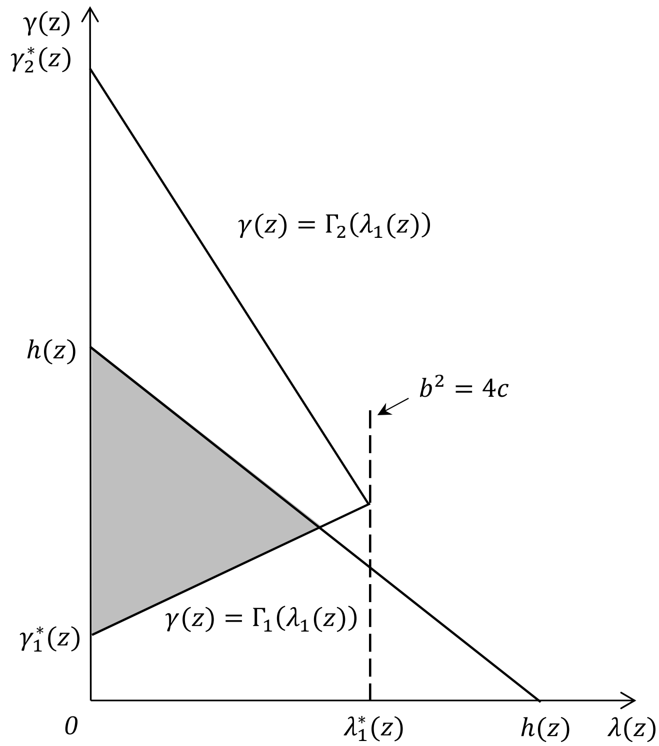

Figure A2: **The set of combinations of $\lambda(z)$ and $\gamma(z)$ that satisfy (A8)**

$$E\left[\Pi^U(z) - \Pi^{dN}(z)\right] = E\left\{\frac{\widetilde{v}(z)}{4h}\left[2h(1+L) - \left(1+\frac{L}{\Phi}\right)\left[\,\widetilde{c}_N(z) + \widetilde{c}_S(z)\right]\right] - \frac{h}{4}\alpha[\Phi-1]^2\right\};$$

$$E\left[\Pi^U(z) - \Pi^{dS}(z)\right] = E\left\{\frac{\Omega}{2}\frac{\widetilde{v}(z)}{4h}\left[2h(1+L) - \left(1+\frac{L}{\Phi}\right)\left[\,\widetilde{c}_N(z) + \widetilde{c}_S(z)\right]\right] + \Omega C - \frac{h}{4}\alpha[\Phi-1]^2\right\}.$$

$E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right] > 0$ iff

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > \frac{\Omega C}{1-\Omega/2} \equiv RHS_1;$$

$E[\Pi^U(z) - \Pi^{dS}(z)] > 0$ iff

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > \frac{E\left\{\frac{h}{4}\alpha\left[\Phi-1\right]^2\right\} - \Omega C}{\Omega/2} \equiv RHS_2;$$

$E[\Pi^U(z) - \Pi^{dN}(z)] > 0$ iff

$$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right] > E\left\{\frac{h}{4}\alpha\left[\Phi-1\right]^2\right\} \equiv RHS_3.$$

Since $\widetilde{c}_N(z) \equiv B(\lambda+\gamma)^z(\beta)^{1-z}$, $\widetilde{c}_S(z) \equiv B\left(\lambda+\gamma w_S\right)^z\left(\beta\right)^{1-z}$ and $\widetilde{v}(z) \equiv \widetilde{c}_N(z) - \widetilde{c}_S(z)$, we have

$E\left[\hat{\Pi}^{dS}(0) - \Pi^{dN}(0)\right] = 0$ and

$$E\left[\hat{\Pi}^{dS}(1) - \Pi^{dN}(1)\right] = E\left\{\frac{\gamma(1-w_S)}{4h}\left[2h(1+L) - \left(1+\frac{L}{\Phi}\right)\left[2\lambda + \gamma(1+w_S)\right]\right]\right\}.$$

$E\left[\hat{\Pi}^{dS}(z) - \Pi^{dN}(z)\right]$ is independent of Ω but increases with z and decreases with w_S.

$RHS_1 \equiv \frac{\Omega C}{1-\Omega/2}$ is independent of z but increases with Ω.

$$\lim_{\Omega \to 0}\left[\frac{\Omega C}{1-\Omega/2}\right] = 0$$

$RHS_2 \equiv \frac{E\left\{\frac{h}{4}\alpha[\Phi-1]^2\right\} - \Omega C}{\Omega/2}$ is independent of z but decreases with Ω.

$$\lim_{\Omega \to 0}\left[\frac{E\left\{\frac{h}{4}\alpha\left[\Phi-1\right]^2\right\} - \Omega C}{\Omega/2}\right] = \infty$$

$RHS_3 \equiv E\left\{\frac{h}{4}\alpha\left[\Phi-1\right]^2\right\}$ is independent of both z and Ω.

Now refer to Figure 4. The lower envelope of $E[\Pi^U(z) - \Pi^{dN}(z)]$ and $E[\Pi^U(z) - \Pi^{dS}(z)]$ intersects the the horizontal axis at $z = \bar{\bar{z}}$.

Suppose $E[\Pi^U(1) - \Pi^{dN}(1)] > 0$ as required in Proposition 2. Thus, the curve of $E[\Pi^U(z) - \Pi^{dN}(z)]$ as a function of z intersects the horizontal axis at an interior z. Suppose we start from $\Omega = \varepsilon$. The curve of $E[\Pi^U(z) - \Pi^{dN}(z)]$ is independent of Ω, and so it is unshifted as Ω increases from ε $(\to 0)$.

$E[\Pi^U(1) - \Pi^{dS}(1)] < 0$ at $\Omega = \varepsilon$. Thus, interior $\bar{\bar{z}}$ does not exist at $\Omega = \varepsilon$. The curve of $E[\Pi^U(z) - \Pi^{dS}(z)]$ shifts up continuously as Ω increases from ε. When Ω increases to $\bar{\bar{\Omega}}$, the curve of $E[\Pi^U(z) - \Pi^{dS}(z)]$ intersects the horizontal axis at $z = 1$. Thus, $\bar{\bar{z}} = 1$ at $\Omega = \bar{\bar{\Omega}}$. As Ω increases from $\bar{\bar{\Omega}}$, the curve of $E[\Pi^U(z) - \Pi^{dS}(z)]$ shifts up. This reduces the value of z at which the curve intersects the horizontal axis. Thus, $\bar{\bar{z}}$ falls from 1 as Ω increases from $\bar{\bar{\Omega}}$.

The curve of $E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right]$ intersects the the horizontal axis at $z = \bar{z}$. When $\Omega = \varepsilon$, $\bar{z} \to 0$, since $E\left[\Pi^{dS}(0) - \Pi^{dN}(0)\right] = 0$. Therefore, when Ω is sufficiently small, $\bar{z} < \bar{\bar{z}}$.

As Ω increases from ε, the curve of $E[\Pi^{dS}(z) - \Pi^{dN}(z)]$ shifts down continuously. This would increase $\bar{z}$. Therefore, as Ω gets sufficiently large $(\Omega = \widetilde{\Omega})$, the following curves all intersect the horizontal axis at $z = \bar{\bar{z}}$: $E[\Pi^U(\bar{\bar{z}}) - \Pi^{dN}(\bar{\bar{z}})] = E[\Pi^U(\bar{\bar{z}}) - \Pi^{dS}(\bar{\bar{z}})] = E[\Pi^{dS}(\bar{\bar{z}}) - \Pi^{dN}(\bar{\bar{z}})] = 0$. Refer to Figure 4.

As Ω increases from $\widetilde{\Omega}$, $\bar{\bar{z}}$ is unchanged, since $\bar{\bar{z}}$ is determined by the intersection of $E[\Pi^U(z) - \Pi^{dN}(z)]$ with the horizontal axis, which is independent of Ω. But the intersection of $E\left[\Pi^{dS}(z) - \Pi^{dN}(z)\right]$

with the horizontal axis continues to rise. Thus, $\bar{z}$ continues to increase as Ω increases from $\widetilde{\Omega}$, until $\bar{z} = 1$ at $\Omega = \bar{\Omega}$.

Thus, there exists a unique $\widetilde{\Omega}$ such that $\bar{\bar{\Omega}} < \widetilde{\Omega} < \bar{\Omega}$ and (i) $\bar{\bar{z}} = \bar{z}$ iff $\Omega = \widetilde{\Omega}$; (ii) $\bar{\bar{z}} < \bar{z}$ iff $\Omega > \widetilde{\Omega}$; and (iii) $\bar{z} < \bar{\bar{z}}$ iff $\Omega < \widetilde{\Omega}$.

PROOF OF PROPOSITION 4

Below we show that $CS_S^U(z) < CS_S^{dN}(z)$ if $\gamma(z) < h(z)\alpha(z)[\Phi(z) - 1]$ for all z.

$CS_S^U(z) < CS_S^{dN}(z)$ iff the following is true:

$$E\left\{[\widetilde{c}_N(z) - \widetilde{c}_S(z)]\left[2h - [\widetilde{c}_N(z) + \widetilde{c}_S(z)]\right]\right\} < E\left\{h\alpha(\Phi - 1)\left[2[h - \widetilde{c}_S(z)] - h\alpha(\Phi - 1)\right]\right\}.$$

This inequality holds if it holds for a certain set of combinations of *ex-post* realizations of $L(z)$, $\Phi(z)$, $h(z)$, $\lambda(z)$, $\beta(z)$, and $\gamma(z)$. This requires that $LHS < RHS$, where

$$LHS \equiv [c_N(z) - c_S(z)]\left[2h(z) - [c_N(z) + c_S(z)]\right];$$

$$RHS \equiv h\,(z)\,\alpha\,(z)\,[\Phi\,(z) - 1]\left[2[h(z) - c_S(z)] - h(z)\alpha(z)[\Phi(z) - 1]\right].$$

First, we note that $LHS = RHS$ when $c_N(z) - c_S(z) = h(z)\alpha(z)[\Phi\,(z) - 1]$. Second, we note that RHS rises monotonically as $h(z)\alpha(z)[\Phi(z) - 1]$ rises, since $q_S^U(z) > 0$ implies $h(z) - c_S(z) - h(z)\alpha(z)[\Phi(z) - 1] > 0$, while LHS is unchanged. Thus, $LHS < RHS$ when $c_N(z) - c_S(z) < h\,(z)\,\alpha(z)[\Phi(z) - 1]$. Since $dE[\widetilde{c}_N(z) - \widetilde{c}_S(z)]/dz > 0$ and $dE[\widetilde{c}_N(z) - \widetilde{c}_S(z)]/dw_S < 0$, the inequality $c_N(z) - c_S(z) < h\,(z)\,\alpha(z)[\Phi(z) - 1]$ holds for any $z \in [0, 1]$ and w_S as long as it holds when $z = 1$ and $w_S = 0$, i.e., iff $\gamma\,(z) < h(z)\alpha(z)[\Phi(z) - 1]$ for all z.