

Universidade de São Paulo  
Faculdade de Filosofia, Letras e Ciências Humanas  
Departamento de Ciência Política  
Programa de Pós-Graduação em Ciência Política

Manoel Galdino Pereira Neto

Determinantes da Adesão a Tratados de Patentes, 1970-2000:  
a Convenção de Paris e o Tratado de Cooperação de Patentes

**São Paulo**  
**2011**

Universidade de São Paulo  
Faculdade de Filosofia, Letras e Ciências Humanas  
Departamento de Ciência Política  
Programa de Pós-Graduação em Ciência Política

**Determinantes da Adesão a Tratados de Patentes, 1970-2000:  
a Convenção de Paris e o Tratado de Cooperação de Patentes**

**Manoel Galdino Pereira Neto**

Tese apresentada ao Programa de  
Pós-Graduação em Ciência  
Política do Departamento de  
Ciência Política da Faculdade de  
Filosofia, Letras e Ciências  
Humanas da Universidade de São  
Paulo para a obtenção do título de  
Doutor em Ciência Política

Orientador: Prof. Dr. Amâncio  
Jorge Oliveira

São Paulo  
2011

Dedico este trabalho à Paula, que me apoiou sempre e suportou pacientemente minhas ausências na frente do computador rodando simulações e escrevendo a tese.

## Agradecimentos

Este trabalho não seria possível sem a ajuda de muitas pessoas. Queria agradecer primeiramente ao meu orientador, Amâncio, que tem me orientado desde o mestrado. A liberdade intelectual, inteligência e visão metodológica dele foram muito importantes e certamente transparecem nesse trabalho.

Queria agradecer também à CAPES e à sociedade brasileira e paulista, que financiou parte dos meus estudos com uma bolsa de doutorado e com a gratuidade da USP, e aos funcionários, particularmente a Rai, e professores e colegas do DCP.

Agradeço também aos professores Mauro Rodrigues, Enivaldo Rocha, George Avelino e Janina Onuki por terem aceito o convite de participarem da banca de defesa da tese de doutorado, e aos professores Fernando Limongi e Naércio Menezes pelos comentários, críticas e sugestões durante a qualificação; e à professora Janina Onuki, que substituiu o prof. Amâncio como presidente da banca de qualificação.

Agradeço ainda ao CAENI e a todos os pesquisadores que por lá passaram, pelo ambiente intelectual, oportunidades e infra-estrutura propiciada. Em especial agradeço ao Umberto por todas as discussões que tivemos, não somente acadêmicas. Certamente não seria hoje um Bayesiano se não fosse pela influência dele. Agradeço ainda ao Rodolpho, Gustavo, Gabriel e Pedro e aos participantes do Seminário de Pesquisa e colegas nos vários cursos que participei.

Por fim, um agradecimento especial à minha mãe, meu pai e meu irmão.

## Resumo

Neste trabalho investigamos os determinantes da adesão de países a dois tratados internacionais de patentes: A Convenção de Paris e o Tratado de Cooperação de Patentes (TCP). Por meio de um modelo hierárquico Bayesiano, apresentamos evidências de que fatores domésticos são importantes para prever adesão aos tratados estudados. Porém, quais fatores são importantes dependem do tipo de tratado. Para o TCP, que é um tratado que visa reduzir custos de transação, a legislação doméstica de patentes não é relevante. Para a Convenção de Paris, que limita as opções de política na área de patente, a legislação doméstica é fator relevante. Nós mostramos também que os ganhos diretos de participar dos tratados, medido pelo número de patentes no exterior, é uma variável importante e positivamente associada à probabilidade de adesão a ambos os acordos. Apresentamos ainda evidências de que variáveis sistêmicas são importantes e que as mudanças no sistema internacional nos últimos 30 anos são fatores importantes para explicar a adesão.

## Palavras-chave:

Patentes, Convenção de Paris, Tratado de Cooperação de Patentes, Ratificação, Modelos Hierárquicos Bayesianos

## Abstract

In this paper we investigate the determinants of the accession of two international patent treaties: the Paris Convention and Patent Cooperation Treaty (PCT). Through a Bayesian hierarchical model, we present evidence that domestic factors are important in predicting accession to the treaties studied. However, what factors are important depends on the type of treaty. For TCP, which is a treaty aimed at reducing transaction costs, the domestic law of patents is not important. For the Paris Convention, which limits the options in the area of patent policy, domestic law is a relevant factor. We also show that the direct gains from participating in treaties, as measured by the number of patents abroad, is an important variable and positively associated with the likelihood of ratification to both agreements. We also present evidence that systemic variables are important and that changes in the international system over the past 30 years are important factors to explain the membership to the treaties.

Key words: Patents, Paris Convention, Patent Cooperation Treaty, Ratification, Hierarchical Bayesian Models

## Sumário

1. Introdução.....	8
2. Revisão da Literatura .....	13
3. Adesão a Tratados de Patentes: Convenção de Paris, Tratado de Cooperação de Patentes e Modelo Teórico.....	22
4. Metodologia Bayesiana.....	44
5. Resultados Empíricos - Convenção de Paris .....	61
5.3. Regime Político.....	65
5.4. Nível Educacional .....	66
5.5. Legislação doméstica em patentes .....	68
5.6. Patentes nos Estados Unidos.....	70
5.7 Investimento Externo Direto (IED).....	71
5.8. Variáveis ao nível do quinquênio.....	72
5.9. Regressão.....	73
5.10. Segundo Nível.....	86
5.11. Ajuste do Modelo .....	88
6. Resultados Empíricos – Tratado de Cooperação de Patentes.....	92
6.1. Métodos e procedimentos empíricos .....	93
6.2. Renda per capita.....	93
6.3. Regime Político.....	95
6.4. Nível Educacional .....	96
6.5. Legislação doméstica em patentes .....	97
6.6. Patentes nos Estados Unidos.....	98
6.7. Investimento Externo Direto (IED).....	99
6.8. Abertura Comercial.....	100
6.9. Variáveis ao nível do quinquênio.....	102
6.10. Modelo de Regressão .....	102

6.11. Segundo Nível.....	109
6.12. Ajuste do Modelo .....	110
7. Discussão dos Resultados e Considerações Finais .....	113
8. Referências Bibliográficas.....	117
9. Anexos .....	126
Anexo A .....	126
Anexo B.....	130

*We need to acknowledge our weaknesses, celebrate our strengths and above all state clearly where we are unsure. It might be bad politics, but it's good science<sup>1</sup>.*

## 1. Introdução

O objetivo do presente trabalho é entender quais os principais determinantes da adesão a tratados internacionais de patentes. Para tanto, estudamos os determinantes da adesão a dois tratados internacionais de patentes, a saber: a Convenção de Paris, tal como revisada em Estocolmo em 1968, e o Tratado de Cooperação de Patentes, firmado em 1968 e que entrou em vigor em 1978.

A Convenção de Paris é um acordo internacional firmado em 1883 sobre propriedade industrial e aborda vários aspectos do direito da propriedade intelectual, como marcas, patentes, indicações de origem e concorrência desleal. Contudo, foi a normatização sobre patentes o aspecto mais importante da Convenção (Gontijo, 2005, Drahos, 1998). O acordo estudado no presente trabalho são os artigos de 1 a 12, do texto revisado em Estocolmo, em 1968, e que tratam justamente dos aspectos substantivos da Convenção. Os artigos 13 a 30 tratam apenas dos aspectos formais e não foram considerados.

O Tratado de Cooperação de Patentes, criado após a Convenção de Paris, busca reduzir custos de transação na obtenção de patentes no exterior. É um tratado importante dentro do regime de patentes, na medida em que é amplamente utilizado pelas empresas e inventores dos estados membro (WIPO, 2009).

Esse trabalho se insere na agenda de pesquisa de regimes internacionais. Tradicionalmente essa literatura tem enfatizado os aspectos sistêmicos e negligenciado os fatores domésticos. Nosso trabalho procura preencher essa lacuna ao estudar tanto fatores domésticos como sistêmicos, com ênfase nos aspectos internos dos estados.

Nesse sentido, seguimos a literatura mais recente sobre adesão e *compliance* a tratados internacionais, embora no presente trabalho concentramo-nos no aspecto da adesão. Segundo essa literatura, o custo de *enforcement* dos tratados, benefícios diretos, conseqüências indiretas (atração de investimento, ajuda estrangeira, ganhos políticos domésticos) e o papel

---

<sup>1</sup> Cameron Neylon.

informativa dos tratados são as principais explicações para a adesão dos países a tratados internacionais.

Vários regimes foram estudados por essa literatura, como direitos humanos, livre-comércio, meio-ambiente e direito trabalhista, mas até o momento não houve estudos sobre o regime internacional de patentes. Assim, a segunda contribuição do trabalho é estender os estudos de adesão para o regime de patentes, área importante no mundo contemporâneo e que até o momento não foi estudado.

De fato, as patentes influenciam diretamente nas capacidades de inovações dos países, difusão e absorção de tecnologia e transferências de recursos entre países por meio do pagamento de royalties e similares.

Nesse sentido, complementamos essa agenda de pesquisa ao estudar dois casos do regime de patentes e qualificamos os achados da literatura ao mostrar que a influência de cada um desses fatores depende do próprio conteúdo do tratado. No caso da Convenção de Paris, que propicia ganhos cooperativos, mas também implica em custos distributivos, custo de *enforcement*, benefícios diretos e ganhos colaterais são relevantes.

Porém, para o Tratado de Cooperação de Patentes, que enfatiza redução de custos de transação e demanda pouco impacto em termos de *enforcement*, os benefícios diretos e consequências indiretas são os fatores mais importantes.

A variável que mede o custo de *enforcement* do tratado é um índice de força da proteção patentária, criado por Ginarte e Park (1997), atualizado posteriormente por Park (2008), mensurado a cada cinco anos para o período estudado. O índice Ginarte-Park foi defasado em cinco anos. As evidências sugerem que para a Convenção de Paris ele está associado positivamente à maior probabilidade de adesão.

Neste estudo, tomamos como *proxy* do benefício direto de um estado com os tratados o número de patentes nos EUA, a renda per capita e o nível educacional do país. Nós mostramos fortes evidências de que quanto maior o número de patentes nos EUA, maior a probabilidade de adesão para ambos os tratados.

Benefícios indiretos ou consequências colaterais não podem ser exauridos, dado que são muitos os possíveis benefícios indiretos do tratado. De todo modo, tomamos como variáveis

que indicam possíveis ganhos indiretos: renda per capita, abertura comercial dos países, o fluxo de investimento externo direto e nível de escolaridade.

Um resultado contra-intuitivo é a associação negativa entre educação e adesão aos tratados. A princípio, esperar-se-ia que países com maior nível de escolaridade aderissem mais a tratados de patentes, pois possuem economia com maior capacidade de inovação e, portanto, de obtenção de patentes. Para o Tratado de Cooperação de Patentes, há uma interação renda-educação e o efeito negativo não é válido para os países desenvolvidos. Para a Convenção de Paris, não há interação educação-renda.

No caso da Convenção de Paris os ganhos diretos para países com alta escolaridade é baixo, posto que têm legislação patentária consolidada e a taxa de inovação pouco seria afetada pelo tratado. Porém, países com baixa escolaridade podem obter ganhos indiretos. Estes vêm no regime internacional uma forma de estimular inovação doméstica e aumentar os retornos à educação dos cidadãos, incentivando-os a investir em capital humano.

No caso do Tratado de Cooperação de Patentes, os ganhos diretos da adesão ao TCP são maiores justamente para países desenvolvidos com nível educacional alto. Porém, para os países pobres com alto nível educacional, o TCP prejudica esses países, pois dificulta a imitação (que eles têm capacidade de fazer devido ao nível educacional). Assim, temos uma explicação para o resultado encontrado: quanto menor a renda, maior o efeito negativo da educação sobre a probabilidade de adesão.

Para a Convenção de Paris, temos fracas evidências de que quanto menor o fluxo de investimento externo direto como proporção do PIB, maior a probabilidade de adesão. Isso sugere que os países com maior fluxo de investimento externo direto teriam pouco a ganhar com a adesão em termos de atração de investimentos, portanto, teriam menor probabilidade de adesão. Temos fracas evidências também de que quanto maior a renda per capita, maior a probabilidade de adesão. Quanto mais desenvolvido o país, mais ele teria a ganhar com a Convenção de Paris. Assim, a taxa de adesão desses países seria mais alta.

Por fim, apresentamos evidências de que há tendências crescentes de adesão no tempo. Esses resultados estão em acordo com a ideia geral de que nos últimas décadas normas visando protegeras inovações com patentes tem ganhado força entre os países, especialmente com o fim da Guerra Fria e consequente vitória ideológica do Capitalismo sobre o socialismo e a criação da OMC em 1994.

O período de análise se estende de 1970 ao ano 2000, e dados quinquenais no formato *Time-Series Cross-Section* (TSCS) foram utilizados em uma amostra com 91 países. Para entender os determinantes de adesão a esses dois tratados, estimamos um modelo hierárquico Bayesiano com efeitos específicos por quinquênio.

Outra contribuição do nosso estudo é metodológica. Várias metodologias diferentes têm sido empregadas para estimação de modelos de adesão a tratados. No entanto, salvo melhor análise, nenhum trabalho empregou modelos hierárquicos Bayesianos para estimar os determinantes da adesão.

Modelos hierárquicos Bayesianos são uma classe de modelos que permitem incorporar na análise estatística a própria estrutura dos dados. Assim, para dados em painel – que é o nosso caso e de quase todos os estudos da literatura-, é possível pensar que os dados são agrupados por países ou por período de tempo. Dados agrupados por países significam que os dados de cada país são correlacionados no tempo, mas não entre países. Dados agrupados por período de tempo significam que há correlação entre países num mesmo período de tempo, mas não entre períodos de tempo.

Esse tipo de modelo na abordagem Bayesiana viabiliza a modelagem dos *clusters*. Dizendo de outro modo, os níveis hierárquicos (efeitos por país ou por tempo, a depender do modelo) podem ser eles mesmos modelados e explicados também por outras variáveis medidas no nível escolhido (por país ou por tempo). Assim, ao contrário de outros trabalhos, no presente trabalho as variáveis sistêmicas são incluídas de maneira mais apropriada no modelo estatístico.

Ainda no campo da contribuição metodológica, nós empregamos *posterior predictive checks* (Gelman e Hill, 2007, Gelman et. al. 2005) para medir os ajustes dos modelos. Esse tipo de checagem é, salvo melhor juízo, inédita na literatura de ciência política no Brasil. *Posterior predictive checks*, similarmente à análise de resíduos, são uma forma de checar o ajuste do modelo contra os dados da amostra, de forma a permitir entender melhor quais aspectos dos dados o modelo não capta e, portanto, inferir em quais aspectos o modelo deve ser melhorado.

O restante desse trabalho está organizado como segue. O capítulo 2 apresenta a revisão da literatura. O capítulo 3 contém uma breve descrição dos dois tratados estudados, nossas escolhas para operacionalização empírica do conceito de adesão e o modelo teórico. O

capítulo 4 apresenta a metodologia Bayesiana. O capítulo 5 apresenta os resultados da estimação dos determinantes da adesão à Convenção de Paris. No capítulo 6 estudamos os determinantes da adesão ao Tratado de Cooperação de Patentes. O Capítulo 7 apresenta as considerações finais.

## 2. Revisão da Literatura

A principal agenda de pesquisa sobre adesão a tratados internacionais é da teoria de Regimes Internacionais, que surgiu cerca de 30 anos atrás (Ruggie, 1975; Keohane & Nye, 1977; Krasner, 1983). Essa vertente teórica preocupa-se entre outros aspectos em explicar porque países cooperam no sistema internacional construindo regimes internacionais. Normalmente, Regime Internacional é definido como consistindo de “principles, norms, rules, and decision-making procedures around which actor converge in a given issue-area” (Krasner, 1982, p. 185), embora alguma controvérsia tenha existido sobre essa definição (Levy et. al., 1995)<sup>2</sup>. De todo modo, parece não haver dúvida de que a teoria de regimes internacionais enfatiza “the institutionalized co-operation of states for managing conflicts and interdependence problems” (Rittberger, 1993, p.9). Como o presente trabalho trata da cooperação internacional, é dentro dessa tradição que o presente trabalho se insere.

No entanto, a teoria de regimes internacionais concentrou-se durante seus primeiros estudos em questões apenas indiretamente relacionadas com nosso problema de pesquisa. De acordo com Mayer, Rittberger e Zürn (1993), a agenda de pesquisa em teoria de regimes se concentrou em entender: i) formação, persistência e desaparecimento dos regimes; ii) categorizar e explicar as propriedades dos regimes (explicar a variação nos tipos de regimes formados); e iii) estudo das conseqüências ou efeitos dos regimes<sup>3</sup>.

Para um estudo sobre os determinantes da adesão a regimes de patentes, importa a literatura sobre formação de regimes, na medida em que teoriza sobre as razões porque estados criariam e, portanto, participariam de um regime. Relacionado com nosso estudo está também a literatura sobre variação nos regimes, já que busca explicar as diferenças de adesão entre um regime e outro e podem, portanto explicar o comportamento que pretendemos estudar. Por fim, a literatura sobre conseqüências dos regimes, que atualmente tem se desenvolvido com foco na questão do *compliance* ajuda a entender os ganhos e custos dos regimes e, portanto, razões pelas quais estados podem aderir ou não a regimes internacionais.

---

<sup>2</sup> Regimes internacionais devem ser considerados como instituições internacionais. Porém, diferenciar-se-iam de organizações internacionais, pois, em última instância, organizações podem agir, regimes não (Hasenclever et. al., 1997, especialmente cap. 2).

<sup>3</sup> De fato, a leitura do livro organizado por Krasner (1983) sobre regimes internacionais discute fundamentalmente essas três questões. Ver também Rittberger (1993), Hasenclever et. al (1997).

Dentro da agenda sobre formação de regimes, a principal corrente é a dos institucionalistas neoliberais (Hasenclever et. al., 1997). Para essa linha de pesquisa, o papel dos regimes é fundamentalmente a troca de informações e cooperação visando evitar resultados subótimos (Keohane, 1989; Keohane, 1984; Axelrod and Keohane, 1986). Em outras palavras, o conteúdo do tratado explicaria a adesão dos países quando tivessem a função principal de reduzir custos de transação e transmissão de informação.

Young (1986, 1989a, 1989b, 1991) acrescenta a essas perspectivas algumas variáveis sistêmicas capazes de aumentar a chance de participação de países em acordos internacionais: ocorrência de um choque exógeno ou crise, como por exemplo a descoberta de um buraco na camada de ozônio; a presença de um líder; e outras variáveis que dizem respeito na verdade ao tipo de acordo, como presença de saliência (acordo claro e familiar aos participantes) e capacidade de *enforcement* do regime. Todas essas características aumentariam a probabilidade de formação de um regime, isto é, dos países aderirem ao regime.

De acordo com Rittberger (1993), a teoria de regimes se desenvolveu também na Alemanha, com uma abordagem distinta da literatura americana. Segundo este autor, surgiu uma literatura baseada na estrutura do problema ou conflito, segundo a qual são as características das questões que determinam se um regime emergirá para administração do conflito. Assim conflitos sobre valores teriam probabilidade muito baixa de formar um regime. Conflitos sobre meios teriam maior probabilidade de formação de regimes que viabilizassem a cooperação. Conflitos poderiam ser também sobre objetos em que os ganhos relativos são mais importantes ou em que os ganhos absolutos é que importam. Nesse último caso, regimes seriam mais fáceis de serem erigidos, ao contrário de conflitos em que predominam ganhos relativos (Efinger, Rittberger e Zürn, 1988, PP. 92-97 *apud* Rittberger, 1993).

Ainda explorando a estrutura do conflito, tipologias baseadas em teoria dos jogos foram propostas. Jogos de coordenação (como a guerra dos sexos) seriam propensos à formação de regimes, pois a cooperação seria pareto superior (Rittberger, 1993). Jogos dilemáticos como o dilema do prisioneiro propiciariam a formação de regimes apenas devido a fatores outros como a sombra do futuro e pequeno número de atores (Rittberger, 1993). Por fim, jogos “Rambo”, ou de *deadlock*, no qual a não-cooperação é a estratégia dominante dificilmente teriam regimes formados.

Entretanto, mais importante para o presente trabalho do que a agenda de pesquisa sobre formação de regimes é a literatura que procura explicar o que Keohane chamou de variação

entre regimes (Keohane, 1993). De acordo com Keohane (1993), a literatura institucionalista dos anos 80 tomou emprestado da microeconomia a metáfora da oferta e demanda para explicar variação na formação e adesão a regimes. Do lado da oferta, a teoria da estabilidade hegemônica propõe que estados hegemônicos têm interesse em criar regimes, servindo como empreendedores políticos e produzindo os bens públicos necessários para promover uma ordem internacional (Kindleberger, 1973). Assim, distribuições mais equitativas do poder estão associadas aos declínios dos regimes (Kindleberger, 1981). Isso se daria porque o país hegemônico teria os maiores ganhos com a criação do regime internacional e, portanto, estaria disposto a tentar resolver o problema da ação coletiva. No que Snidal (1985) chamou de versão benevolente da TEH, o país hegemônico é o responsável pela produção do regime. Já na versão coercitiva, o país hegemônico coage outros países a produzir conjuntamente o bem público ou regime (Gilpin, 1987).

Porém, vários autores criticaram a teoria da estabilidade hegemônica, pois argumentam que muitas vezes os países líderes são antes obstáculos que promotores da ordem internacional, ou que não são os responsáveis pela adesão dos demais países ao regime internacional em questão, como teria sido o caso do padrão-ouro (Fiori, 2005; Walter, 1993; Strange, 1987).

Do lado da demanda, Keohane (1993) argumenta que é necessário analisar os determinantes das preferências dos estados. Instituições domésticas e/ou coalizões políticas dominantes podem alterar as preferências dos estados sobre participar ou não de regimes internacionais. Outro argumento avançado por Keohane (1993) é que quanto maior a interdependência, maior a propensão para participar de regimes internacionais. Ademais, Keohane (1993) argumenta que regimes que obtêm sucesso se tornam mais benéficos e podem atrair mais países.

Outra teoria onde podemos encontrar explicações sobre os comportamentos de países no que tange à adesão ou não a regimes internacionais é a teoria das potências médias ou países intermediários. Essa teoria iniciou-se a partir da categorização de Keohane (1969) dos tipos de Estados entre: *system-determining*; *system-influencing*; *system-affecting* e *system-inneffectual*. O primeiro tipo seriam os Estados determinantes ou líderes, possivelmente um estado hegemônico, e são capazes de mudar o sistema internacional unilateralmente; o segundo tipo de estados são aqueles influentes, mas não hegemônicos, capazes de influenciar o sistema, mas não de dominá-lo; o terceiro tipo seriam os estados intermediários, que

poderiam influenciar o sistema desde que agissem coletivamente e preferencialmente em arenas multilaterais; por fim, o último tipo seria incapaz de influenciar o sistema.

A partir dessa classificação, desenvolveu-se uma literatura sobre países intermediários, supondo-se sempre certa peculiaridade no comportamento desses países<sup>4</sup>. Possivelmente devido à própria dificuldade metodológica em definir quem são os países intermediários, não há estudos sistemáticos a respeito do comportamento dos países intermediários. De todo modo, parte significativa da literatura sugere um comportamento multilateralista e coletivista dos países intermediários (Cooper, et. al., 1993), o que sugeriria a princípio maior probabilidade de adesão aos regimes internacionais.

Porém, os países intermediários não teriam um único comportamento, já que podem, por exemplo, escolher formar coalizões contra-hegemônicas. Lima (2005) aponta que o Brasil enquanto país intermediário atuaria ora como *system-affecting state*, buscando influenciar o sistema internacional; ora como potência emergente e portanto buscando se adequar (e integrar) à ordem internacional. Assim, dependendo os benefícios da adesão a um tratado, o país intermediário poderia se comportar de um modo ou de outro.

O problema dessas teorias, contudo, é a difícil operacionalização empírica das principais variáveis determinantes do comportamento internacional dos países. Além disso, ao enfatizar variáveis sistêmicas, negligenciaram o papel dos fatores domésticos nos determinantes da decisão dos estados em participar de tratados internacionais. De acordo com Zürn, a desimportância dada aos fatores domésticos deriva, pelo menos em parte, da crença segundo a qual “*systemic explanations are generally superior to subsystemic ones because of their parsimony*” (Zürn, 1993, p. 12). Junne (1990, *apud* Zürn, 1993) argumenta que a ênfase em fatores domésticos desacreditaria a noção de que a ação de estados seria governada por instituições internacionais.

Mais recentemente, a literatura de regimes internacionais avançou nessa agenda de pesquisa sobre variáveis domésticas dos estados, concentrando esforços em estudar fatores específicos que explicam a ratificação e *compliance* a tratados. Assim, a literatura adicionou mais especificidade às argumentações gerais, pouco específicas, da teoria de regimes.

Alguns estudos fundamentalmente empíricos mostraram que em alguns casos democracias eram mais propensas à cooperação internacional (Gaubatz, 1996). Alguns autores

---

4 cf. Sennes (2001).

argumentaram então que uma das razões porque democracias se engajariam mais em compromissos internacionais é para dificultar a mudança de políticas por governos futuros (Moravcsik, 2000, Reinhardt, 2002). Ademais, a literatura sugere que novos regimes teriam incentivos em aderir a tratados, já que desejam ser aceitos na ordem internacional e a adesão a tratados é uma forma de atingir esse objetivo (Hathaway, 2007).

Não obstante, um *puzzle* empírico demandava explicações. Democracias sistematicamente violavam os compromissos assumidos em alguns tipos de tratados, especialmente sobre direitos humanos. Além disso, havia muitas democracias que não aderiam a alguns tratados, enquanto autocracias o faziam. Em outras palavras, era necessário explicar por um lado por quais razões estados adeririam a tratados se não pretendiam cumpri-los, e por outro lado por quais motivos estados com menos “violações” ao conteúdo dos tratados não aderiam a eles.

Assim, autores como Hathaway (2007) retomam o argumento mais geral da teoria de regimes segundo a qual países aderem a tratados se os benefícios superam os custos. Porém, para além das questões sistêmicas, introduzem a importância dos regimes políticos domésticos como fatores determinantes dos benefícios e custos de um tratado. Os custos de *não-compliance* seriam diferentes entre ditaduras e democracias. Em regimes autoritários, a “regra da lei” seria mais frouxa e, portanto, o custo de *não-compliance* menor. Já para democracias, o custo de *não-compliance* seria maior. Assim, como democracias e ditaduras enfrentariam custos diferentes para adesão a tratados, a decisão de aderir a eles seria também diferente a depender do regime político.

Vreeland (2008) argumenta que no caso de tratados de direitos humanos, ditaduras com algum grau de competição política experimentam mais dissidência e, portanto, mais tortura. Porém, ditaduras são obrigadas a fazer concessões aos opositores e acabam por aderir a tratados de direitos humanos, como a Convenção Contra Tortura. Vale destacar que, aqui, os benefícios da adesão são não apenas sistêmicos, mas envolvem ganhos na política doméstica do país.

Tomz (2002), entretanto, argumenta que nem sempre democracias enfrentam custos maiores de *não-compliance* de compromissos internacionais. Analisando a moratória da Argentina de 2001, ele argumenta que a opinião pública apoiava a moratória e, portanto, a violação das regras internacionais. Assim, o efeito da democracia sobre custos de *compliance* dependeria das preferências da opinião pública.

Hathaway (2007) argumenta que as conseqüências colaterais de adesão a tratados – para além das conseqüências legais de *enforcement* – são importantes determinantes da decisão dos países de aderirem a tratados internacionais. De acordo com Hathaway,

*These collateral consequences arise from the reactions of domestic or transnational actors to the state's decision to commit to the treaty. Unlike legal enforcement, collateral consequences are not structured by the terms of the treaty but instead arise from interactions that fall outside its bounds. Because they are not part of the formal structure of the treaty, these effects are often ignored. Yet they can prove to be just as important as, if not more important than, formal legal enforcement of the treaty requirements in influencing states' behavior* (2007, p. 595).

Assim, adesão a tratados podem provocar reações que podem gerar custos ou benefícios. A depender os custos e benefícios específicos gerados por cada tratado, países podem ser mais ou menos propensos a aderir os tratados.

Mais recentemente, a literatura, utilizando a teoria de agente-principal, argumenta que um papel importante dos tratados está na sinalização para audiência doméstica ou internacional (Farber, 2002; Simmons 2000; Mansfield e Pevehouse, 2006). Se aderir a um tratado é custoso, então a adesão do estado pode ser um sinal informativo para audiência doméstica e externa sobre as intenções do líder político. O tipo de sinal emitido dependeria das instituições domésticas (democracia ou autocracia) e o tratado em questão.

No caso de tratados de livre-comércio, políticos em democracias utilizariam tratados internacionais para sinalizar ao eleitor que estão adotando boas políticas. A previsão de tais modelos, corroboradas empiricamente, é que regimes democráticos seriam mais propensos a se engajar em tratados de livre-comércio (Mansfield et al., 2000); Dai, 2002; Mansfield et al., 2002; Mansfield e Reinhardt, 2008). Em tratados de direitos humanos, ditadores sinalizariam para a audiência que eles não temem ser processados por violações de direitos humanos, pois pretendem permanecer no poder (Hollyer e Rosendorf, 2011).

Essa teorização é importante por reforçar o papel dos tratados em transmitir informações críveis sobre os tipos dos atores. Em geral, a *rationale* para transmissão de informação está no tipo de regime do país. Assim, em democracias, tratados transmitem algum tipo de informação, e em autocracias tratados transmitem outro tipo de informação.

Além de fatores domésticos, a literatura também tem testado empiricamente o papel de variáveis sistêmicas na explicação da adesão dos países a tratados internacionais,

particularmente em estudos de tratados de meio-ambiente. Fatores como grau de participação de terceiros no tratado e efeitos de rede foram incorporados nas análises.

Bernauer, Kalbhen, Koubi e Spilker (2010) estudam o papel de fatores externos *vis-a-vis* fatores domésticos na decisão de ratificação de 180 países em 225 tratados multilaterais de meio-ambiente durante o período 1950-2000. A unidade de análise é país-tratado por ano, e o país é removido da base de dados assim que ratifica o tratado. Efeitos específicos por país e tratado são incluídos na análise. Os autores concluem que efeitos externos como participação em organizações internacionais, intensidade comercial e percentual de países que ratificaram o tratado, são mais importantes que fatores domésticos (renda e democracia).

Baccini e Koenig-Archibugi (2010) estudam a adesão a convenções da Organização Internacional do Trabalho (OIT). Os autores argumentam que estados buscam congelar padrões trabalhistas e evitar a utilização de padrões trabalhistas inferiores como arma competitiva e que aspectos normativos da comunidade internacional influenciam na adesão a normas trabalhistas. Em que pese uma engenhosa estratégia de identificação de efeitos de socialização entre estados, é possível argumentar que o estudo não diferencia causalmente os efeitos estimados entre homofilia (preferências similares), contágio e efeitos de ambiente comum<sup>5</sup>.

Frank (1999) estuda tratados ambientais e, por meio de um modelo de equações estruturais com variáveis latentes, argumenta que o principal preditor da adesão a esses tratados é o grau de conectividade dos estados com a sociedade mundial na área ambiental. Mais uma vez, vale a ressalva sobre a dificuldade em diferenciar homofilia, contágio e ambiente comum.

Como se vê, há vários estudos sobre adesão a tratados internacionais nos mais diversos temas. Todavia, não encontramos estudos sobre adesão a tratados de propriedade intelectual em geral, ou tratados de patentes em particular. Assim, passamos agora a revisar alguns estudos sobre economia política das patentes que, embora não estudem adesão a tratados, podem jogar luz sobre nosso problema de pesquisa.

Do ponto de vista doméstico, se o governo está interessado em maximizar o bem-estar doméstico, então o que deve determinar os incentivos para adoção de proteção intelectual mais forte diz respeito à variação marginal na inovação tecnológica em virtude de

---

<sup>5</sup> Sobre as dificuldades de identificar o papel da homofilia, contágio e ambiente compartilhado em estudos de rede, ver Lyons (2011) e Shalizi e Thomas (2011).

fortalecimento do regime doméstico de proteção à propriedade intelectual. Porém, quando considerado do ponto de vista internacional, temos uma situação similar a um dilema do prisioneiro, na medida em que se assemelha à provisão de um bem público, gerando o clássico problema de ação coletiva (Olson, 1965). Na ausência de qualquer coordenação entre os países, o conjunto de proteção doméstica à propriedade intelectual de todos os países tende a ser subótimo.

Grossman & Lai (2004) desenvolveram um modelo de teoria dos jogos para explicar a provisão de proteção internacional como um bem público global. Os autores supõem dois países que diferem em suas capacidades de inovação e tamanho de mercado. Eles derivam duas previsões teóricas: i) quanto maior o mercado de um país, mais o país ganha com leis de patentes; ii) quanto maior a capacidade de inovação de um estado, mais ele se beneficia com leis patentárias mais fortes.

Neste modelo, para fins de simplificação, temos dois grupos de países – do Norte e do Sul. Se um país do Norte fortalece suas leis de propriedade intelectual doméstica, então isso diminui a importância do mercado do Sul para os lucros de um inovador presente em países do Norte e do Sul. Logo, reduz a responsividade do regime de patentes doméstico do país do Sul. Nesse caso, o país do Sul, sabendo que seu mercado tem menos importância para um inovador, é estimulado a diminuir a proteção (doméstica) à propriedade intelectual. Ademais, devido ao tamanho do mercado do norte ser maior e devido a maiores capacidades de produzir P&D, maior é a responsividade do inovador a leis de patentes no Norte que no Sul. Assim, criam-se interesses conflituosos entre os países e uma tendência à produção subótima da proteção à propriedade intelectual no nível mundial, qualquer que seja o nível ótimo.

Esses resultados teóricos são consistentes tanto com estudos qualitativos sobre as políticas de proteção à propriedade intelectual dos países atualmente desenvolvidos quando ainda não eram industrializados (Shang, 2004), como em estudos quantitativos mais recentes que sugerem correlação entre capacidade de P&D e direitos de patentes (Park, 2008).

Assim, as variáveis domésticas fundamentais a determinar os incentivos de um país a fortalecer o seu regime doméstico são i) a capacidade de inovar relacionada *vis-a-vis* o de outros países, ii) tamanho de mercado do país *vis-a-vis* o de outros países e iii) nível de proteção de propriedade intelectual nos outros países.

Assim, se os países têm incentivos à cooperação internacional para a produção do bem público e alcançar alguma forma de superar o problema do carona, um regime internacional de propriedade intelectual que preveja tratamento igual entre os países é mais difícil ainda de ser produzido, na medida em que os ganhos não são iguais entre os países. Assim, mesmo que o problema da ação coletiva esteja superado, não é óbvio porque países aceitariam participar de regimes de patentes com proteção mínima.

Outro estudo indiretamente relacionado com a nossa pesquisa é o trabalho de Chen e Puttitanum (2005). Os autores estudam a associação entre legislação doméstica de patentes e desenvolvimento econômico. Eles concluem, similarmente a outros estudos (Falvey, Foster e Greenaway, 2006), que a relação é em formato de U. Assim, nos extremos (países ricos e países pobres) teríamos comportamento similar, e no nível intermediário (países de renda média) teríamos outro padrão. Os autores concluem países mais desenvolvidos e menos desenvolvidos tem legislação com proteção patentária mais forte, e países com renda média teriam menor nível de proteção. Segundo Chen e Puttitanum (2005), economias de renda média, caracterizadas por muitas firmas imitadoras e poucas inovadoras se beneficiariam com um nível de proteção patentária fraco. Países mais desenvolvidos teriam mais empresas inovadoras e, portanto, mais a ganhar com um maior nível de proteção de patentes. E países menos desenvolvidos, sem um complexo industrial com capacidade imitativa ou inovadora, prefeririam conceder forte proteção patentária para atender a interesses de empresas multinacionais.

Esses estudos mostram a importância do nível de desenvolvimento econômico e capacidade industrial na formação dos interesses dos estados na área de legislação de patentes. Embora esses trabalhos se concentrem na arena doméstica, é possível estender alguns dos argumentos para a arena internacional, objeto do presente trabalho.

A literatura de regimes internacionais e a literatura sobre *compliance* também podem fornecer subsídios relevantes para a formulação teórica do presente trabalho. Segundo essa literatura, os custos legais, conseqüências colaterais e o aspecto informacional são os fatores-chave na determinação dos custos e benefícios de adesão a um tratado.

### **3. Adesão a Tratados de Patentes: Convenção de Paris, Tratado de Cooperação de Patentes e Modelo Teórico**

O Presente Capítulo apresenta os principais aspectos dos dois tratados escolhidos para análise e desenvolve o modelo teórico. O objetivo é não somente resumir as principais disposições que podem influenciar a decisão dos países de aderir a cada tratado, mas mostrar como operacionalizamos nossa variável dependente: adesão a tratados de patentes. Assim, inicialmente apresentamos um resumo do conteúdo da Convenção de Paris e do Tratado de Cooperação de Patentes. Em seguida, mostramos como codificamos empiricamente a adesão dos países aos respectivos tratados. Por fim, apresentamos o modelo teórico que guiará a análise estatística.

#### **Convenção de Paris**

A Convenção de Paris foi criada em 20 de março de 1883, em Paris, para proteção da propriedade industrial. Os países membros da convenção constituem a União para a proteção da propriedade industrial.

A Convenção teve várias revisões ao longo do tempo, a saber: Bruxelas em 4 de dezembro de 1900, em Washington a 02 de julho de 1911, em Hague a 06 de novembro de 1925, em Londres a 02 de junho de 1934, em Lisboa a 31 de outubro de 1958 e em Estocolmo a 14 de julho de 1967.

A Convenção, tal como revisada em Estocolmo, entrou em vigor em 26 de abril de 1970 em relação aos artigos 13 a 30, e a 16 de maio no que tange aos artigos 1 a 12.

Os aspectos substantivos da Convenção de Paris estão contidos nos artigos 1 a 12, e os aspectos procedimentais nos artigos subsequentes (13 a 30 na Convenção de Paris). A partir da revisão de Estocolmo, em 1967, facultou-se aos países a adesão ao tratado – por meio de assinatura seguida de ratificação ou acesso – separadamente desses dois conjuntos de artigos. Em outras palavras, um Estado pode aderir apenas aos artigos 1 a 12, apenas aos artigos 13 a 30, ou à totalidade dos artigos de uma única vez.

O artigo 1 da Convenção de Paris especifica os temas objetos de normatização do tratado: patentes, modelos de utilidades, desenhos industriais, marcas e nomes comerciais, marcas de serviços, indicações de fonte ou origem e repressão de concorrência desleal.

Dentre esses temas, o mais importante para nós diz respeito às normas sobre patentes. A razão para tal é que os demais aspectos da propriedade industrial regulada pela Convenção são menos controversos do que os de patentes. Além disso, é a proteção patentária objeto de maior conflito de interesse entre os países no que tange à adesão ao tratado. Assim, nesse trabalho trataremos indistintamente a propriedade industrial e a proteção de patentes, exceto quando especificamente mencionada a diferença.

Do ponto de vista dos princípios que guiam a Convenção de Paris, podemos destacar o tratamento nacional, o direito de prioridade e o princípio da independência das patentes (WIPO, 2009; Gontijo, 2005):

## 1. TRATAMENTO NACIONAL

O artigo 2 da Convenção de Paris, afirma:

*Nationals of any country of the Union shall, as regards the protection of industrial property enjoy in all the other countries of the Union the advantages that their respective laws now grant, or may hereafter grant, to nationals; all without prejudice to the rights specially provided for by this Convention. Consequently, they shall have the same protection as the latter, and the same legal remedy against any infringement of their rights, provided the conditions and formalities imposed upon nationals are complied with” (WIPO, 2008, p. 2)<sup>6</sup>*

Esse princípio é importante porque garante que as legislações patentárias dos países não podem beneficiar apenas os residentes. Durante muito tempo, principalmente no século XIX era comum que patentes fossem concedidas apenas para nacionais ou residentes, discriminando contra invenções no estrangeiro. Assim, se de fato patentes estimulam inovação, essa discriminação contra o estrangeiro não ajudava a resolver o problema da produção do bem público global. Com o aumento da interdependência comercial ao longo do século XX, esse princípio é importante para garantir que as empresas possam obter patentes nos vários países onde comercializam seus produtos.

Além disso, e ao contrário do *Trade Related Aspects of Intellectual Property* (TRIPs) da Organização Mundial do Comércio (OMC), nenhuma cláusula de reciprocidade é requerida para aplicação do princípio de tratamento nacional. Isto significa que, se a legislação de um país A concede 20 anos de proteção patentária para invenções e um país B concede 10 anos de proteção, então A não pode conceder 20 anos de proteção apenas para seus residentes sob

---

<sup>6</sup> A tradução em português, tal como redigida pelo decreto N. 75.572, de 08 de abril de 1975, contém a tradução oficial, pelo governo brasileiro, do texto da Convenção de Paris.

a alegação de que não há reciprocidade no país B<sup>7</sup>. De fato, a Convenção de Paris não contém muitas disposições visando a harmonizar a legislação dos países no que tange à forma como a proteção patentária deveria ocorrer.

## 2. DIREITO DE PRIORIDADE E PRINCÍPIO DE INDEPENDÊNCIA DE PATENTES

Conforme vimos na origem da Convenção de Paris, ela surgiu, entre outros motivos, como uma resposta à questão prática da disseminação de informações de invenções entre países, especialmente nas Feiras Mundiais que ocorriam no século XIX. O direito de prioridade, expresso no artigo 4, é uma dessas respostas a essa necessidade prática, na medida em que concede a um requerente de uma patente um período de prioridade no qual ele pode requisitar um pedido de patente em outro Estado membro da União de Propriedade Industrial.

Como argumenta Gontijo, o objetivo do direito de prioridade é:

*... evitar apropriação indevida de informações incluídas nos pedidos de patente e, ao mesmo tempo impedir conflitos em casos de dois ou mais inventos sobre o mesmo objeto, decidiu-se assegurar àquele que tenha feito o pedido de patente em um dos países da União um prazo de prioridade (que hoje é de 12 meses) para realizar o depósito em outros países, durante o qual nenhum outro pedido invalidará o seu, nem qualquer publicação ou exploração do invento (Gontijo, 2005, p.5).*

O prazo de prioridade é de 12 meses para patentes e 6 meses para modelos de utilidade. Modelos de utilidade são invenções menores, em geral resultado de aperfeiçoamento de ferramentas e equipamentos existentes, que garantem proteção similar a patentes, porém em geral por período mais curto (WIPO, 2011). Nem todos os países garantem proteção para modelos de utilidade<sup>8</sup>. Na prática, pode ser considerado como uma proteção patentária mínima global em sentido fraco. Isso porque seria possível exibir e mesmo comercializar um produto em países sem a patente, desde que o pedido já tivesse sido feito em um dos países membro da União. Se uma empresa copiasse o produto dentro do prazo de 12 meses, o inventor original poderia entrar com um pedido de patente e, pelo princípio de prioridade,

---

<sup>7</sup> Essa é uma possível razão para a demanda dos países com proteção patentária mais extensa de harmonização da legislação patentária no mundo num nível mínimo, como aconteceu com o TRIPs.

<sup>8</sup> Para uma lista de países que atualmente garantem proteção de modelos de utilidades, ver esta nota no website da OMPI: [http://www.wipo.int/sme/en/ip\\_business/utility\\_models/where.htm](http://www.wipo.int/sme/en/ip_business/utility_models/where.htm) (acessado em junho de 2011).

obtê-la e, portanto, tornar uma violação da patente a cópia do concorrente. É contudo uma proteção em sentido fraco porque, em primeiro lugar, a concessão de uma patente num país A não obriga um país B a também conceder a patente; em segundo lugar, porque em países em que não há proteção patentária pela legislação para o produto em questão o inventor nada pode fazer.

O fato de que a concessão (ou negação) de uma patente num país A não vincula de maneira alguma a concessão num país B foi consagrado pela Convenção de Paris e figura como **princípio da independência das patentes**. Esse princípio é importante para garantir a soberania dos países no que tange aos procedimentos administrativos e legislativos dos países em matéria de propriedade industrial. Por outro lado, o tempo de proteção patentária de uma inovação não pode se basear na data de registro de pedido em outro país, já que as patentes são independentes.

Pensando em termos de princípios gerais, esses são os aspectos mais importantes da Convenção de Paris para o presente estudo. Como se sabe, o ato de Estocolmo foi a última revisão da Convenção de Paris. Membros antigos do tratado não eram obrigados a aceitar esta última revisão. Entretanto, novos membros não podem aderir a versões anteriores da Convenção e devem aceitar o ato de Estocolmo para se tornarem membros da União.

Do ponto de vista substantivo, duas revisões dos últimos atos em comparação com o tratado original são importantes. Em primeiro lugar, tornou-se mais difícil a emissão de uma licença compulsória (Gontijo, 2005). Como se sabe, uma licença compulsória é a autorização, por parte de um país, de que outras empresas produzam um produto patenteado ou utilizem um processo de produção patenteado, mediante remuneração para o dono da patente. Em outras palavras, a licença compulsória é um mecanismo para suspensão temporária do monopólio garantido pela patente. A Convenção de Paris autoriza a emissão de licença compulsória para tentar obrigar o dono da patente a fabricar a inovação dentro do país, sendo designado esse fato por falta de exploração local da patente. Do mesmo modo, os países podiam determinar a caducidade da patente por falta de exploração local da mesma.

Como argumenta Gontijo (2005, pp. 6-7),

*a ameaça contida na revogação da patente por meio da caducidade, (...) é uma das formas de extinção da patente. Decretada a extinção, a patente cai em domínio público, o que permite a qualquer pessoa ou empresa explorar o objeto da invenção. (...) A caducidade é um instrumento poderoso, porque sua aplicação é automática após o esgotamento do prazo para a exploração*

*local, desde que comprovada a não exploração local. Por isso mesmo, viu-se lentamente substituída nas revisões da Convenção pela figura da licença compulsória.*

Posteriormente, as sucessivas revisões dificultaram a adoção desses mecanismos por partes dos países. Dentre as principais modificações, está a exigência de que a licença compulsória seja emitida de forma exclusiva, isto é, apenas para uma única empresa, para produção doméstica e após negociações com a empresa terem falhado. Além disso, não poderá ser emitida se o titular da patente apresentar razões legítimas para a não exploração local da patente.

Textualmente, diz o artigo 4, inciso 4, do ato de Estocolmo:

*Não poderá ser pedida licença obrigatória, com o fundamento de falta ou insuficiência de exploração, antes de expirar o prazo de quatro anos a contar da apresentação do patente, ou de três anos a contar da concessão da patente, devendo aplicar-se o prazo mais longo; a licença será recusada se o titular da patente justificar a sua inação por razões legítimas. Tal licença obrigatória será não-exclusiva só será transferível, mesmo sob a forma de concessão de sublicença, com a parte da empresa ou do estabelecimento comercial que a explore (decreto n. 75.572, de 08 de abril de 1975).*

Tudo bem pesado, o ato de Estocolmo aumentou a proteção padrão de patentes em países membros da União. Não por outra razão, vários países, como o Brasil, que eram membros da Convenção desde sua criação se recusaram a aceitar as novas revisões até os anos 90 (Gontijo, 2005). No Brasil vigorava a Convenção tal como revisada em Haia, 1925, no que diz respeito aos artigos 1 a 12.

### **O Tratado de Cooperação de Patentes**

Dentro do sistema da Convenção de Paris, o registro de uma patente em vários países tem de ser feito no período de 12 meses para que o titular da invenção possa reivindicar o direito de prioridade. Isto significa que no período de um ano é necessário traduzir o pedido de patente, contratar advogados especializados nos vários países e pagar as taxas para pedido de patentes nos países. Tudo isso sem ter a certeza de que a patente será válida ou que poderá recuperar o investimento com a mesma.

Para evitar então os custos de transação associados à existência de vários escritórios de patentes nos países, em 1956 o comitê executivo da União para a Propriedade Industrial requisitou ao BIRPI (Birô Internacional Reunido para a Propriedade Industrial), predecessor

da atual OMPI, um estudo para reduzir os custos duplicados tanto dos requerentes quanto dos escritórios de patentes (WIPO, 2009, p. 277). Como consequência, textos foram preparados e em junho de 1970, uma conferência em Washington adotou um tratado, designado por Tratado de Cooperação de Patentes (TCP). O tratado entrou em vigor em 24 de janeiro de 1974 e tornou-se operacional em primeiro de junho de 1978 (WIPO, 2009, p. 277).

Vale destacar que o TCP não funciona como um substituto para os escritórios de patentes, nem tampouco tem a função de conceder ou negar um pedido de patente. O que o TCP faz é, tão somente, viabilizar a redução dos custos de transação associados ao pedido de patentes em vários escritórios. Esses pedidos são então divididos em duas fases, a nacional e internacional. Na fase internacional, o trabalho é quase todo feito pelo TCP ou agências designadas pelo TCP. Na fase nacional, os escritórios de patentes avaliam os pedidos segundo seus procedimentos usuais.

De acordo com a própria OMPI (WIPO, 2009, pp. 277-278), o TCP:

*establishes an international system which enables the filing, with a single Patent Office (the "receiving Office"), of a single application (the "international application") in one language having effect in each of the countries party to the PCT which the applicant names ("designates") in his application;*

*- provides for the formal examination of the international application by a single Patent Office, the receiving Office;*

*- subjects each international application to an international search which results in a report citing the relevant prior art (mainly published patent documents relating to previous inventions) which may have to be taken into account in deciding whether the invention is patentable;*

*- provides for centralized international publication of international applications with the related international search reports, as well as their communication to the designated Offices;*

*- provides an option for an international preliminary examination of the international application, which gives the applicant and subsequently the Offices that have to decide whether or not to grant a patent, a report containing an opinion as to whether the claimed invention meets certain international criteria for patentability.*

Como se pode perceber pelo resumo apresentado acima, o TCP não interfere nas decisões dos escritórios de patentes. Porém, efetivamente reduz os custos de transação, ao reduzir a necessidade ou os custos de pesquisa sobre *prior art*, emite um parecer sobre o atendimento

de alguns critérios, disponibiliza o pedido de patente para todos os escritórios designados pelo requerente e facilita a pesquisa dos pedidos de patentes por qualquer pessoa ou entidade.

Contudo, o TCP interfere sobre as legislações dos países, na medida em que a forma e conteúdo requerido num pedido de patente internacional (isto é, utilizando o TCP) deve ser o mesmo em todos os estados membros. De fato, afirma o Handbook da própria OMPI: “No national law may require compliance with requirements relating to the form or contents of the international application different from or additional to those which are provided for by the PCT.”(WIPO, 2009, p. 279).

Com relação ao pagamento de taxas, paga-se uma taxa única numa mesma moeda, no escritório que recebe o pedido. Essas taxas devem cobrir os custos do escritório que recebe os pedidos e os custos do Birô Internacional. As taxas dos vários escritórios nacionais são pagos apenas em data futura e residentes de países menos desenvolvidos tem direito a pagar taxas menores que as usuais.

Após a fase internacional, o requerente da patente pode analisar os relatórios emitidos e a busca por *prior art* e decidir se pretende prosseguir com o pedido de patente nos vários escritórios de patentes ao redor do mundo. Caso desista formalmente ou simplesmente não declare o desejo de dar prosseguimento ao processo na fase nacional, o pedido caduca e não é necessário pagar as taxas dos escritórios de patentes nacionais. Dessa forma, ele evitou vários custos que incorreria se operasse pelo procedimento normal da Convenção de Paris com 12 meses para pedidos de prioridade.

Para se ter uma idéia da utilização do pedido de patente por meio do sistema criado pelo TCP, de 1978 a 2003, quase 900 mil pedidos internacionais de patentes foram registrados no Birô Internacional da OMPI. Apenas em 2003 foram recebidos mais de 100 mil pedidos de patentes, o que equivaleria a mais de sete milhões de pedidos nos escritórios nacionais de patentes (WIPO, 2009, p. 5, anexo). Como se vê por esses números, o TCP efetivamente reduz os custos de transação no plano internacional.

### **Classificação dos Tratados**

Como vimos, há diferenças importantes entre a Convenção de Paris e o TCP. Para os propósitos desse trabalho, podemos classificar os tratados basicamente em dois tipos. Há tratados que buscam a redução nos custos de transação para obtenção de proteção patentária nos diversos países, e há aqueles que buscam aumentar a proteção patentária no interior dos

países, modificando a legislação doméstica destes. Em linhas gerais podemos designar o Tratado de Cooperação de Patentes como sendo do primeiro tipo, isto é, visa primordialmente a reduzir custos de transação, enquanto a Convenção de Paris é mais bem classificada como sendo do segundo tipo, isto é, visa a aumentar a proteção patentária doméstica dos países signatários por meio de modificação na legislação.

Baseado na literatura (Rittberger, 1993) que sugere que o tipo de problema que um regime e/ou tratado procura resolver é importante, podemos classificar a Convenção de Paris como um tratado sobre valores, posto que aborda a limitação da política industrial por meio da emissão de licenças compulsórias e proteção global de patentes. E o TCP pode ser classificado como um tratado sobre meios, isto é, como reduzir custos de transação por meio de um procedimento único de registro de patentes.

Outra possibilidade de classificação, ainda com base na literatura, é considerar a classificação dos tipos de jogos. Nesse caso, a Convenção de Paris seria um jogo dilemático. Isso porque o princípio do tratamento nacional impõe aos países com maior nível de proteção de patentes maior contribuição para o bem público global, que seria a proteção global de patentes. E o TCP seria um jogo de coordenação, em que o mais importante é a adoção de um padrão comum.

Essas diferenças são importantes porque influenciam nos custos e benefícios de adesão ao tratado. Na seção em que discutimos o modelo teórico nós retomamos essa distinção, para argumentar que o tipo de tratado influencia no efeito esperado de cada uma das variáveis preditoras.

### **Da Adesão a Tratados de Patentes**

O objetivo do presente trabalho é estudar a adesão de países a tratados de patentes. Conceitualmente, a adesão a um tratado por parte de um Estado ocorre quando o Estado se compromete a cumprir as determinações do tratado por meio da internalização do conteúdo do tratado na legislação do país. Contudo, como sói acontecer na passagem da teoria para a operacionalização empírica de conceitos, a classificação dos casos no interior do conceito de adesão nem sempre é clara. A presente seção discute os critérios adotados no trabalho para operacionalização empírica do conceito de adesão.

Nos tratados sob a égide da OMPI, há três categorias possíveis para adesão: assinatura do tratado, ratificação do tratado e acesso (*accession*).

A diferença entre ratificação e assinatura é que na ratificação o tratado foi internalizado à norma interna do país signatário, enquanto que na assinatura o país ainda não internalizou o tratado e, portanto, suas disposições legais não estão em vigor. A ratificação, portanto, ocorre após a assinatura do tratado e significa a internalização da norma no corpo legal do país, após a assinatura.

O acesso a um tratado é similar à ratificação. Porém, podemos dizer que o acesso é a ratificação sem ser precedida da assinatura, isto é, no acesso tudo se passa como se a assinatura e ratificação se dessem simultaneamente.

A lógica para essa diferenciação é que muitos tratados não estão abertos indefinidamente para a assinatura. Nos tratados sob a égide da Organização Mundial de Propriedade Intelectual, é comum que, após um tempo, não seja mais permitida a assinatura do tratado sem ratificação. Nesses casos, apenas o acesso é permitido. Dizendo de outro modo, após expirado esse prazo, tudo se passa como se a assinatura do tratado fosse feita em conjunto com a ratificação (simultânea) e isso configura o acesso (*accession*). Nesses casos, não se diz que um país assinou e ratificou o tratado, mas apenas aceitou a ele.

Como ilustração dessas diferenças, vejamos o que diz o artigo 62 do Tratado de Cooperação de Patentes (WIPO, 2010):

(1) *Any State member of the International Union for the Protection of Industrial Property [Convenção de Paris] may become party to this Treaty by:*

- (i) *signature followed by the deposit of an instrument of ratification, or*
- (ii) *deposit of an instrument of accession.*

O artigo 62 é explícito em afirmar que um Estado se torna parte vinculada ao tratado apenas por meio da assinatura e depósito do instrumento de ratificação ou depósito do instrumento de acesso. Porém, os países não podem simplesmente escolher a forma de aderir ao tratado, isto é, por meio de assinatura seguida de ratificação ou acesso. Diz o artigo 67 do mesmo tratado, sobre assinatura e idiomas, no parágrafo 2 (WIPO, 2010): “*This Treaty shall remain open for signature at Washington until December 31, 1970*”.

Ou seja, após 1970, a adesão por meio de assinatura seguida da ratificação não seria possível. Apenas o mecanismo do acesso estaria disponível para um Estado interessado se tornar parte do tratado.

Além disso, a própria entrada em vigor do tratado pode ter disposições transitórias, condicionais ao cumprimento de certos requisitos mínimos em termos de adesão.

As negociações para adoção do Tratado de Cooperação de Patentes (TCP) foram concluídas em 1970. Isso significa que a partir de 1970 estava aberta aos países membros da Convenção de Paris a assinatura do Tratado. Porém, pelo acordo inicial, o tratado só teria validade jurídica internacional a partir do momento que um número mínimo de países ratificasse o tratado. Segundo o artigo 63 do referido tratado:

#### Artigo 63.

*(1)(a) Subject to the provisions of paragraph (3), this Treaty shall enter into force three months after eight States have deposited their instruments of ratification or accession, provided that at least four of those States each fulfill any of the following conditions:*

*(i) the number of applications filed in the State has exceeded 40,000 according to the most recent annual statistics published by the International Bureau,*

*(ii) the nationals or residents of the State have filed at least 1,000 applications in one foreign country according to the most recent annual statistics published by the International Bureau,*

*(iii) the national Office of the State has received at least 10,000 applications from nationals or residents of foreign countries according to the most recent annual statistics published by the International Bureau.*

*(b) For the purposes of this paragraph, the term "applications" does not include applications for utility models.*

*(2) Subject to the provisions of paragraph (3), any State which does not become party to this Treaty upon entry into force under paragraph (1) shall become bound by this Treaty three months after the date on which such State has deposited its instrument of ratification or accession.*

*(3) The provisions of Chapter II and the corresponding provisions of the Regulations annexed to this Treaty shall become applicable, however, only on the date on which three States each of which fulfill at least one of the three requirements specified in paragraph (1) have become party to this Treaty without declaring, as provided in Article 64(1), that they do not intend to be bound by the provisions of Chapter II. That date shall not, however, be prior to that of the initial entry into force under paragraph (1).*

Como se percebe pela leitura do artigo 63, havia um critério de relevância de adesão para que o tratado entrasse em vigor, já que o termo “aplicações” trata de pedidos de registro de patentes. Para ser ter uma ideia do que representavam na década de 1970 os requisitos de

registro de pedidos de patentes em outros países, apresentamos o número de patentes registradas nos EUA por residentes de outros países:

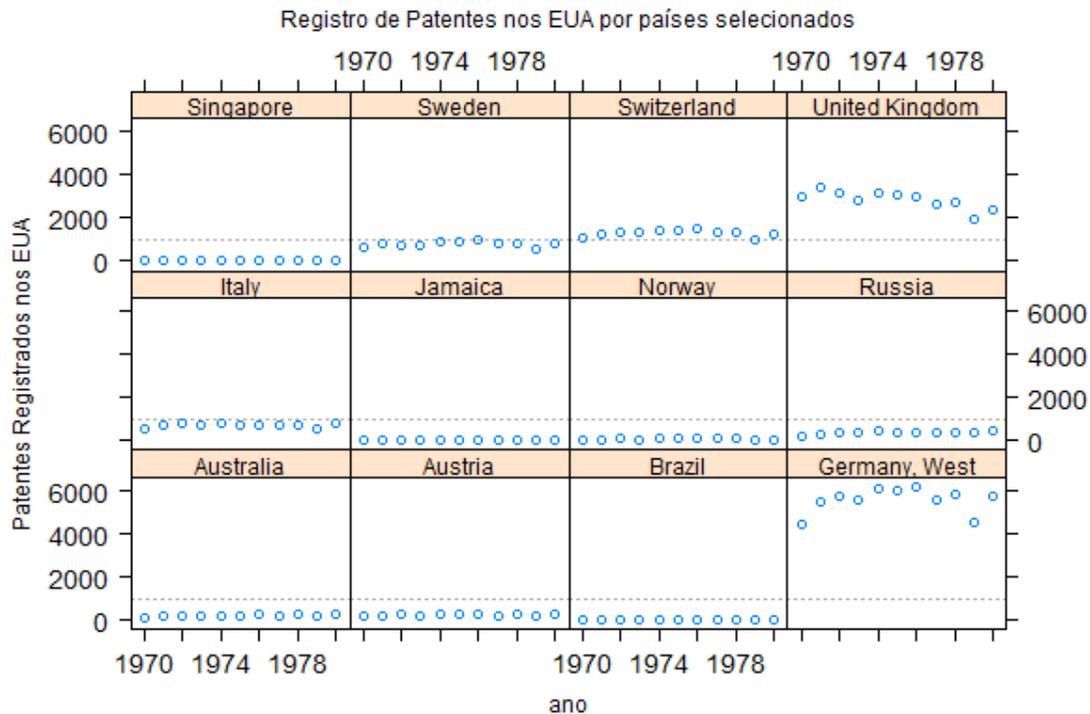


Gráfico 3.1 O gráfico mostra o registro de patentes nos Estados Unidos em países selecionados durante os anos 70. A linha pontilhada representa o número de 10.000 patentes registradas, número mínimo para atender ao critério iii do artigo 63.

Como se pode observar pelo gráfico 3.1 acima, dentre esses países selecionados, apenas Alemanha Ocidental, Reino Unido, Itália, Noruega, Rússia (URSS), Suíça e Suécia tinham mais de 1000 pedidos de patentes (ou por volta de 1000 patentes) nos EUA. Como naquele momento os EUA, alguns países da Europa e Japão eram os principais países em que se buscava pedir patentes, é uma boa Proxy para o pedido de registro de patentes em âmbito global os pedidos nos EUA. Isso significa que praticamente apenas países desenvolvidos cumpriam todos os requisitos dos países cuja adesão faria o tratado entrar em vigor. Contudo, apenas quatro países precisavam preencher os requisitos mínimos, o que na prática significava quatro países desenvolvidos mais uma grande potência – para ultrapassar o pedido de 40.000 registros patentes no próprio país. De fato, o critério mais exigente era a necessidade de o Estado ter recebido pelo menos 40.000 pedidos de patentes por ano. Na década de 70, os países que conseguiam atender a esses critérios, segundo dados da própria OMPI, eram:

Estados Unidos, Reino Unido, Japão, URSS, Alemanha e França (para alguns anos apenas).

Portanto, não é simples determinar em que ano um dado país aderiu a um tratado e nem se assinatura, ratificação e acesso têm o mesmo significado. Apenas para fins de ilustração, três perguntas se sobressaem como mais complicadas: 1) o que melhor retrata a data de adesão: a data de assinatura, a data de ratificação ou data de entrada em vigor?; 2) a decisão de um país aceder ao tratado é diferente do país que assinou e ratificou o tratado ao mesmo tempo? 3) um país que não acedeu a um tratado é equivalente a um que assinou, mas não ratificou?

Para os fins do presente trabalho, tomamos a data de ratificação ou data de acesso ao tratado como data de adesão. Informações sobre a assinatura do tratado serão desconsideradas da análise. Assim, quando nos referirmos à adesão ao tratado, isso significa ratificação se este instrumento foi utilizado ou acesso ao tratado se a ratificação não era possível.

Outro problema para definição da data de adesão ocorre para aqueles países que: 1) eram colônia de um país parte do tratado; 2) se tornaram entidades independentes como resultado da subdivisão de Estados previamente existentes e partes do tratado; 3) resultado da fusão, unificação ou incorporação de outros países.

No primeiro caso, decidimos, seguindo a própria OMPI, não considerar ex-colônias como partes vinculadas ao tratado apenas porque o país colonizador era parte do tratado e as obrigações do colonizador se estendiam por todas as colônias. Na verdade, países só entram na base de dados a partir do momento em que passam a existir como Estados independentes e, portanto, apenas a data de adesão dos Estados independentes é levada em consideração.

O segundo caso, que trata do surgimento de países a partir do desmembramento ou dissolução de outro Estado existente é mais complicado. Nesses casos, a OMPI considera a data de adesão de um país resultado do desmembramento de outro país membro do tratado como a data de declaração de independência do novo país. Em outras palavras, se um país A é parte do tratado no instante  $t$ , e há uma separação do país em  $t+1$ , com surgimento de um novo Estado B, a OMPI considera que esse novo Estado B “herda” o compromisso anterior.

Porém, esse novo país pode fazer ou não uma declaração de continuação de aplicação (ou de validade) do tratado, em data posterior. A declaração de continuação de aplicação/validade significa que o novo país notifica legalmente a OMPI de que ele está vinculado legalmente ao tratado. Dizendo de outro modo, essa declaração é o reconhecimento oficial do novo Estado

de que ele aceita “herdar” as obrigações legais do tratado assinado pelo Estado anterior do qual ele era parte.

Não é o nosso objetivo aqui discutir, do ponto de vista jurídico essa questão. Não somente nos falta competência para tanto, como o mais importante para nós é definir qual data reflete melhor o conceito de adesão que estamos empregando. Para explicitar o critério adotado, iremos utilizar alguns exemplos ilustrativos.

Caso exemplar nessa questão é a Eslováquia, tornada independente em 1993. Na página na internet no sítio da OMPI sobre o Tratado de Cooperação de Patentes, a Eslováquia consta com data de adesão de 1991, referente à data de adesão da antiga Tchecoslováquia. Porém, a data de declaração de continuação de aplicação/validade é de 1993, e a entrada em vigor do tratado no mesmo ano. Nesse caso, temos de definir se iremos codificar a data de adesão apenas em 1991 ou 1993.

Outro exemplo é a Sérvia. Esta aparece como tendo assinado o TCP em 1970 (referente à Yugoslávia, já que a Sérvia não existia como entidade independente na época), ratificação em 1996 (referente à Federação da Yugoslávia) e entrada em vigor em 1997. Porém, a declaração de continuação foi realizada apenas em 2006.

O critério adotado nesses casos foi tomar a data de declaração de continuação como data da adesão, se a declaração existir. Caso inexista, tomamos a data de entrada em vigor do tratado como data de adesão.

O terceiro caso, sobre unificação de países, é ilustrado pela união da Alemanha Oriental com a Alemanha Ocidental. Em se tratando de unificação, anexação ou reunificação de países, decidimos considerar cada estado como uma entidade independente previamente à unificação e após a unificação, apenas o novo Estado como entidade independente. Na prática, isso significa que, por exemplo, tanto a Alemanha Oriental quanto Alemanha Ocidental estão na base de dados como países distintos, porém são retiradas da base de dados após a unificação. Assim, temos dados da Alemanha Oriental e Ocidental entre 1970 até 1989. Após 1990, apenas o país Alemanha passa a ter dados na base.

Outra questão importante é o que fazer com os países que “perdem” parte do seu território como resultado de movimentos separatistas, mas continuam a existir como entidade independente após essa separação. O caso mais ilustrativo é o da União Soviética, que após a separação das repúblicas socialistas soviéticas nos anos 90, passa a ser designada por Rússia.

No presente trabalho, decidimos considerar esses países como indistintos, isto é, para todos os efeitos, a Rússia é o mesmo país que a União Soviética na nossa base de dados.

Por fim, os países podem ratificar o tratado com algumas reservas, isto é, sem aceitar integralmente todas as partes do tratado. Cada tratado explicita quais artigos, cláusulas e parágrafos podem ser objeto de reservas. Para nós, contudo, o problema é que isso cria uma adesão desigual. Ratificar com reservas pode não ser a mesma coisa que ratificar sem reservas.

Assim, se a reserva diz respeito a aspectos substantivos do tratado, nós separamos a adesão do tratado por partes ou conjunto de artigos. Em outras palavras, subdividimos o tratado em partes iguais às quais os países podem aderir ou não. Se a reserva diz respeito apenas a questões jurídicas da norma internacional de tratados ou procedimentais do tratado, não fazemos distinção entre adesão com e sem ressalvas.

No caso do TCP isso não é relevante e consideramos a adesão integral do tratado. No caso da Convenção de Paris, tal como revisada pelo ato de Estocolmo em 1967, separamos o tratado em dois conjuntos de artigos: artigos 1 a 12 e artigos 13 a 30. Assim, é possível que um país adira aos artigos 1 a 12, mas não 13 a 30; adira aos artigos 13 a 30, mas não 1 a 12; ou ainda que adira a ambos os artigos, sejam em datas diferentes ou não. Portanto, na prática tudo se passa para nós como se os dois conjuntos de artigos fossem tratados diferentes objeto da decisão de um país aderir ou não a cada parte.

Os artigos 1 a 12 são a parte mais substantiva da Convenção de Paris. Os artigos 13 a 30 cuidam apenas da parte procedimental, sobre como emendar o tratado e especificação da contribuição financeira dos países, entre outros aspectos. No presente trabalho estudamos apenas a adesão aos artigos 1 a 12 da Convenção de Paris, pois o que queremos entender é quais fatores explicam a adesão a tratados que versam sobre patentes. Se estudássemos a adesão aos artigos 13 a 30, estudaríamos a adesão aos aspectos formais do tratado, e esse não é nosso objetivo.

### **3.1. O Modelo Teórico**

Em geral, a análise estatística dos dados segue mais ou menos o seguinte *script*: o pesquisador formula alguma hipótese sobre o efeito de A em B, coleta dados e então testa a hipótese nula de que o efeito é zero. Se o p-valor for menor do que 5%, então se conclui que efeito de A sobre B é significativo (Shrodt, 2010; Hubbard & Lindsay, 2008; Gill, 1999).

Além disso, acreditamos que raramente a prática acadêmica se coaduna com os preceitos do ritual tradicional. Assim, ao enunciar nossas hipóteses, não imaginamos que o efeito delas possa ser zero. Sua função é tão somente organizar as idéias sob as quais erezimos o modelo estatístico a ser estimado.

### **3.2. Da utilidade do modelo**

Em nossa perspectiva, o objetivo de todo modelo (seja ele puramente formal ou estatístico) é ser útil. É difícil definir formalmente o que faz um modelo mais ou menos útil, e também, operacionalizar empiricamente uma medida de utilidade de um modelo. Contudo, em linhas gerais, podemos dizer que um modelo é útil se ele permite entender a realidade estudada. Além disso, acreditamos que todos os modelos são falsos, já que suas características contradizem aspectos da realidade de alguma maneira.

Nesse sentido, a construção de um modelo pode ser feita de várias maneiras. Porém, em última instância, acreditamos que é necessário que envolva tanto considerações teóricas quanto o teste empírico da teoria, num processo de ida e volta. Em outras palavras, um modelo inicial é formulado teoricamente, em seguida dados são coletados e analisados – tanto por meio de estatísticas descritivas quanto por meio de estimação de modelos baseados nas teorias – e a teoria é revisada, levando a novos dados e novos modelos a serem estimados, até que o modelo seja considerado adequado aos dados. Assim, o modelo final estimado é uma mistura de considerações teóricas e modificações induzidas pela comparação do ajuste do modelo aos dados (Gelman, Meng e Stern, 1996; Gelman, 2003; Gelman, 2004). Grosso modo, modelos mais bem ajustados – como quer que se meça ajuste aos dados – são modelos mais úteis.

Não pretendemos que essa perspectiva de filosofia da ciência sobre a prática científica seja original nem consensual<sup>9</sup>. Mas como não é nosso objetivo discutir essas questões a fundo, mas tão somente apresentar a perspectiva de onde trabalhamos, convidamos o leitor a consultar o trabalho recente de Gelman e Shalizi (2010) para uma exposição detalhada e rigorosa com a qual concordamos<sup>10</sup>. O importante aqui é, portanto, explicar o objetivo e

---

<sup>9</sup> A respeito do aspecto controverso de algumas dessas idéias, veja o artigo de Gelman (2008) e comentários de outros autores na mesma edição do *journal Bayesian Analysis* em que o autor argumentou de um ponto de vista de um suposto anti-bayesiano.

<sup>10</sup> Na verdade, é mais preciso dizer que fomos profundamente influenciado pelos trabalhos de Gelman em termos de metodologia estatística e filosofia da ciência, de forma que nossas idéias refletem em grande medida o que Gelman tem escrito, inclusive no blog do autor.

justificativa para o presente capítulo. Nós apresentamos as razões teóricas para a estimação do modelo estatístico. O modelo final e sua interpretação, porém, envolvem não somente as considerações teóricas, mas também a comparação com os dados da amostra e o posterior refinamento que adveio dessa comparação.

### 3.3. Argumentos Teóricos

Nesta seção nós desenvolvemos os argumentos teóricos e nossas hipóteses de pesquisa para cada uma das variáveis incluídas. O objetivo de formular hipóteses é apresentar o sentido esperado das variáveis sobre a probabilidade de adesão. Dentro da agenda de pesquisa sobre *compliance*, as hipóteses são formuladas segundo os benefícios diretos, conseqüências indiretas/colaterais e custos de *enforcement* para cada tratado.

Num nível mais abstrato, temos o seguinte argumento: i) quanto maior o benefício direto de um tratado, tudo o mais constante, maior a probabilidade de adesão; ii) quanto maior o benefício indireto/colateral, maior a probabilidade de adesão; iii) quanto maior o custo direto – de *enforcement* –, menor a probabilidade de adesão e; iv) quanto maior o custo indireto/colateral, menos a probabilidade de adesão.

#### Educação e renda per capita

Numa economia capitalista sem patentes, conforme argumentou Smith (2011), uma inovação tecnológica que não seja apenas um bem final implica para a firma inovadora a melhoria na produtividade do trabalho sem, contudo, significar aumento na quantidade nem intensidade do capital. Com o passar do tempo, outras firmas iriam copiar a inovação e aumentar a produtividade delas. No agregado, haveria um aumento da produtividade da economia. Abramovitz (1956) e Solow (1957) enfatizaram a mudança tecnológica como principal fator explicativo do crescimento econômico. De fato, tornou-se convencional afirmar que as diferenças de desenvolvimento entre os países devem-se fundamentalmente a diferenças na **Produtividade Total dos Fatores** (PTF), e não nos insumos capital e trabalho. Como se sabe, se pensarmos em termos de uma função de produção agregada,  $f(k, l) = y$ , a diferença entre o produto  $y$ , e os insumos  $(k, l)$ , é justamente o resíduo designado por PTF. Contudo, não é fácil determinar o quanto da PTF deve-se à tecnologia, instituições ou cultura.

De todo modo, numa economia com proteção patentária, a mesma inovação seria protegida por uma patente. Isso significa não somente que a imitação e difusão da tecnologia são dificultadas, mas que a empresa pode vender como mercadoria a inovação. Em outras

palavras, absorção de tecnologia se torna custosa para quem a adquire e lucrativa para quem a vende.

Do ponto de vista econômico, como se sabe, boa parte da discussão normativa sobre os ganhos de bem-estar proporcionados por um sistema de patentes está em avaliar o *trade-off* entre ganhos de longo prazo e prejuízos de curto prazo (Romer, 1990; Huger, Moore e Snyder, 2002; Maskus, 2000; Boldrin e Levine, 2007). Supostamente, no longo prazo teríamos mais inovações num sistema com patentes do que num sistema sem patentes, e no curto prazo maior dificuldade de difusão das inovações e aumento no custo da aquisição de novas tecnologias.

Essa assimetria entre ganhadores e perdedores numa economia com e sem patentes é também refletida em termos globais, isto é, em se tratando de economias no interior de estados nacionais (McCalman, 2001; Diwan e Rodrik, 1991). Países com economias sofisticadas produtoras de inovação tendem a se beneficiar mais da introdução de proteção patentária que economias mais atrasadas. Estas últimas seriam consumidoras de inovação tecnológica, mas não produtoras, de forma que isso implicaria na transferência de recursos desses países de economias atrasadas para economias mais avançadas (Kanwar, 2010; Acemoglu, Aghion e Zilibotti, 2002).

No caso do TCP, que é um tratado que facilita a obtenção de patente no exterior, temos um benefício direto para os países mais desenvolvidos. Assim, países desenvolvidos devem ter alta taxa de participação. Por países desenvolvidos entendemos países com alta renda per capita e alto nível de escolaridade.

Países de renda mais baixa e nível educacional maior tendem a se prejudicar com tratados como o TCP, na medida em que não são produtores de inovação, mas tem capacidade de imitação. Assim, o TCP é um custo direto para esses países.

Países de renda baixa e nível educacional baixo consomem pouca tecnologia estrangeira e, portanto, tem baixo custo direto de adesão ao TCP.

Por fim, países de renda alta e baixa escolaridade consomem bastante tecnologia na forma de produtos finais para os consumidores, mas tem baixa capacidade de absorção de tecnologia que aumenta a produtividade da economia. Assim, tem baixo custo direto e benefício indireto para os consumidores.

Com base nessa análise de custos e benefícios, podemos formular as seguintes hipóteses:

1. Quanto maior a renda per capita, maior a probabilidade de adesão, tudo o mais constante
2. Condicional à renda per capita alta, nível de escolaridade não deve influenciar na probabilidade de adesão
3. Condicional à renda per capita baixa, quanto menor o nível de escolaridade, maior a probabilidade de adesão.

No caso da adesão à Convenção de Paris os ganhos diretos para países com alta escolaridade são baixos, posto que a taxa de inovação doméstica é pouco afetada pelo tratado. Porém, países com baixa escolaridade podem obter ganhos indiretos. Estes vêm no regime internacional uma forma de estimular inovação doméstica e aumentar os retornos à educação dos cidadãos, incentivando-os a investir em capital humano. Assim, podemos formular a hipótese:

1. Quanto menor o nível de escolaridade, maior a probabilidade de adesão à Convenção de Paris, tudo o mais constante.

Com relação à renda per capita, países mais ricos tendem a ter mais empresas transnacionais e, portanto, tendem a se beneficiar mais da cláusula de preferência nacional contida na Convenção de Paris (Braga & Fink, 1998; Lee & Mansfield, 1996). Por outro lado, países pobres tem mais dificuldade de utilizar licenças compulsórias como instrumentos de política industrial e, portanto, maior custo com o tratado. Assim, podemos formular a hipótese:

2. Quanto maior a renda per capita, menor a probabilidade de adesão, tudo o mais constante.

### **Patentes no Exterior**

Outro fator possivelmente relevante para nós é a obtenção de patentes no exterior. A obtenção de uma patente no exterior por parte de uma empresa significa que ela pretende obter o monopólio sobre a exploração da inovação em outros mercados além do mercado doméstico. Para essas empresas, tanto a redução nos custos de transação na obtenção de patentes em terceiros países quanto a garantia de tratamento nacional em terceiros países é benéfica para elas. Assim, países com maior número de patentes no exterior tenderão a se beneficiar da participação dos tratados. Formalmente, temos:

**Tudo o mais constante, quanto maior a demanda por patentes no exterior de uma economia de um país, maior a probabilidade do mesmo participar dos tratados.**

Quanto à operacionalização empírica da demanda por patentes, decidimos utilizar as patentes obtidas nos EUA como *proxy*. Uma vez que os EUA são a partir do pós guerra a principal economia do mundo, as patentes obtidas nesse país são uma aproximação razoável para a demanda de proteção no exterior por parte dos demais países (exceto é claro os próprios EUA<sup>11</sup>). Essa relação deve ser tanto mais forte no caso do TCP, que visa a reduzir os custos de transação de obtenção de patentes no exterior.

### **Legislação Doméstica**

A literatura de adesão a regimes tem enfatizado o papel da legislação doméstica na decisão dos países aderirem ou não tratados. Quanto maior a probabilidade de *compliance* por parte de um país, maior a probabilidade de adesão. Nesse sentido, países com legislação mais favorável à concessão de patentes tenderiam a participar da Convenção de Paris, já que a possibilidade de *compliance* é maior, além do custo de modificar a legislação doméstica seria menor.

A operacionalização dessa variável se dá por meio do índice Ginarte- Park. Em um artigo de 1997, Ginarte e Park (1997) criaram um índice de direitos de patentes para 110 países, quinquenalmente, para o período 1960-1990, posteriormente atualizado até 2005 (Park, 2008). O índice é resultado de uma pontuação atribuída aos países de acordo com a satisfação de vários critérios. Quantos mais critérios satisfeitos, mais pontos o país recebe. Cinco aspectos das leis de patentes dos países foram definidos como critérios a serem analisados: i) extensão (setorial) da cobertura de patentes; ii) participação em tratados internacionais; iii) provisões para perda de proteção; iv) mecanismos de *enforcement*; v) duração da proteção. Em cada categoria foram analisados vários fatores e, para cada fator, uma pontuação é atribuída se o estado atender aos critérios estabelecidos em cada um dos fatores. O máximo valor por categoria é 1, e o mínimo valor é 0. O índice é resultado da soma não ponderada de cada categoria, de forma que o valor mínimo do índice é zero e o valor máximo é cinco. Quanto maior o valor do índice, maior a “força” da proteção patentária. Os detalhes da construção do índice podem ser encontrados nos artigos supracitados.

---

<sup>11</sup> Embora para o caso dos EUA nós deveríamos a rigor utilizar outra medida, não cremos que teremos tanto problema em utilizar as patentes obtidas pelos próprios EUA nos EUA, já que são o país com maior pedido de patentes tanto nos EUA quanto no resto do mundo.

Contudo, o índice não pode ser utilizado diretamente por nós, na medida em que a participação em dois tratados internacionais, a Convenção de Paris e o TCP, são alguns dos critérios utilizados pelos autores para criação do índice. O índice portanto contém como componentes as nossas variáveis dependentes.

Decidimos então diminuir do índice final os valores adicionados no índice original pela participação nos dois tratados, de forma a evitar o problema de que qualquer associação encontrada pudesse derivar do fato de que um componente do índice é a própria variável dependente. Assim, por exemplo, os EUA em 1995 tinham um índice de 4,88. Porém, a participação no TCP e na Convenção de Paris contribuíram cada um com 0,20 no índice, de forma que subtraindo esses valores, temos que os EUA teriam um índice de 4,44. Nesse caso, o índice não tem mais como valor máximo 5, mas 4,6. Nós poderíamos reescalonar o índice, aumentando o peso da participação em outros tratados (como o TRIPs) para que mantivesse valor máximo 5. Porém, como queremos enfatizar os aspectos domésticos com essa variável, decidimos não reescalonar as variáveis. Além disso, não temos acesso aos componentes individuais por país, mas tão somente ao valor agregado do índice, de forma que não seria possível o reescalonamento.

Formalmente, a relação entre legislação doméstica e adesão a tratados de patentes pode ser expressa como segue:

**Tudo o mais constante, quanto mais forte a proteção da legislação de patentes de um país, maior a probabilidade do mesmo participar dos tratados.**

#### **Investimento Externo Direto (IED)**

As relações entre as variáveis anteriores e nossa variável dependente foram construídas a partir da ótica dos interesses de países mais desenvolvidos em participar dos tratados. Quais seriam então os interesses dos países menos desenvolvidos em participar desses tratados?

Países menos desenvolvidos, mesmo sem capacidade de inovação, precisam adquirir tecnologia. Um benefício colateral da adesão a tratados de Patentes é que os países utilizem a proteção patentária para atrair investimento estrangeiro. De fato, a literatura tem mostrado que esse efeito é real (Braga & Fink, 1998; Lai, 1998, *apud* Hu e Jaffe, 2007; Lee & Mansfield, 1996).

Assim, tratados internacionais funcionam como sinalizadores para a comunidade internacional de que as regras institucionais não irão mudar. Essa sinalização é tanto mais crível no caso da Convenção de Paris, porque permite distinguir claramente o que é uma política industrial aceitável de uma política industrial arbitrária que emite licenças compulsórias sem critérios reconhecidos internacionalmente. Assim, o efeito deve ser mais forte para a Convenção de Paris.

Porém, países com alto nível de investimento externo direto tem pouco a ganhar com a adesão, já que são países com instituições consolidadas e pouco arriscadas para o investidor estrangeiro. Assim, são os países com menor nível de investimento externo direto que tendem a obter os maiores benefícios colaterais. Formalmente, nossa hipótese então é:

Tudo o mais constante, quanto menor o fluxo de investimento externo direto, maior a probabilidade de adesão à Convenção de Paris.

### **Abertura Comercial**

Com o aumento da interdependência econômica entre os países, a obtenção de proteção patentária no exterior torna-se cada vez mais importantes para as empresas que exportam e/ou produzem essas mercadorias para consumo no mercado mundial. Assim, países com alto grau de participação no comércio mundial relativo ao tamanho da sua economia tenderão a participar de tratados de patentes como a Convenção de Paris e o TCP.

Parte da literatura sugere, por exemplo, que países mais atrasados tecnologicamente, mas incorporados à economia global, teriam incentivos a assinar tratados de patentes (Hu e Jaffe, 2007; Falvey, Foster e Greenaway, 2006; Chen e Puttitanum, 2005; Gould e Gruben, 1996). Embora os países sem proteção patentária pudessem legalmente adotar tecnologias sem custo de *royalties* e afins, a exportação de produtos com essas tecnologias estariam proibidas para os países que protegessem essas tecnologias com patentes. Nesse caso, o melhor que esses países fariam seria justamente participar dos tratados.

Temos portanto um benefício direto para os países mais abertos comercialmente participarem dos tratados, especialmente o PCT, que reduz os custos de transação.

Por outro lado, países menos integrados podem utilizar o PCT para reduzir os custos de transação de ampliar o fluxo de comércio de mercadorias patenteadas. Assim, os benefícios indiretos são positivos para os países com menor abertura comercial.

Não temos portanto, uma direção esperada *a priori* para a associação entre abertura comercial e adesão.

### **Regime Político**

Conforme argumentado na revisão da literatura, os estudos de adesão a tratados tem enfatizado a questão da sinalização no interior de regimes políticos resultante da adesão a tratados. De um lado, democracias são mais propensas a aderir a tratados internacionais de comércio e transparência, por exemplo, pois a adesão serve como sinalização aos eleitores de que o líder está adotando políticas corretas. Além disso, são uma forma de dar estabilidade às políticas, isto é, são estratégias de *tying hands* contra eventuais líderes populistas.

De outro lado, parte da literatura argumenta que ditadores que pretendem não cumprir tratados de direitos humanos adeririam a ele para sinalizar para a audiência doméstica de que vão fazer o necessário para permanecer no poder, incluindo tortura que viola tratados internacionais assumidos pelo país.

Em ambos os casos, tratados servem como sinalização crível para uma audiência doméstica. O que o tratado sinaliza depende tanto do conteúdo do tratado quanto o tipo de regime político do país. No caso de patentes, a segunda história claramente não faz muito sentido. Ditaduras que pretendem violar tratados de patentes não irão assinar o tratado para sinalizar que pretendem permanecer no poder.

Porém, é possível que a adesão a tratados de patentes, especialmente no caso da Convenção de Paris, seja uma sinalização de que o país está seguindo políticas corretas, tanto para eleitores quanto para grupos de interesse. Nesse sentido, democracias deveriam ser mais propensas a aderir à Convenção de Paris. No caso do TCP, não esperamos que o regime político faça alguma grande diferença, exceto pelo argumento mais geral de que democracias seriam mais cooperativas no plano internacional. Formalmente, podemos enunciar essa expectativa como:

**Tudo o mais constante, países democráticos são mais prováveis de aderir a tratados de patentes do que ditaduras.**

## 4. Metodologia Bayesiana

O objetivo central do presente capítulo é apresentar a metodologia estatística utilizada no presente trabalho, a saber: modelos hierárquicos Bayesianos. Como contudo a própria inferência Bayesiana é pouco conhecida na ciência política brasileira, faremos inicialmente uma introdução à modelagem Bayesiana, para que os termos e conceitos utilizados na análise estatística no próximo capítulo possam ser compreendidos mais facilmente pelo leitor não familiarizado com a abordagem Bayesiana.

### 4.1. O Processo de Geração dos Dados

A adesão de países a tratados internacionais é uma decisão que possivelmente envolve muitos fatores. Instituições domésticas, perfil econômico e impacto do tratado sobre a legislação doméstica são algumas das variáveis mais óbvias. De todo modo, a decisão é uma variável binária, que assume apenas dois valores: adesão ou não adesão. Parece fazer sentido, portanto, modelar esse processo de decisão por meio de um modelo de escolha discreta, como logit ou probit. Numa das várias formas de se interpretar o modelo logit<sup>12</sup>, assume-se que existe uma variável latente,  $z_{i,t}$ , que representa a utilidade para um país  $i$  no ano  $t$  de ser parte do tratado. Se a utilidade de participar do tratado for maior do que a utilidade de não participar, então o país escolhe fazer parte do tratado. Caso contrário, permanece não adere a ele. Essa utilidade é uma função linear de um conjunto de preditores, tal que:  $z_{i,t} = X_{i,t} \beta + \varepsilon_{i,t}$ , com erros  $\varepsilon_{i,t}$  independentes e com distribuição logística. Porém, a utilidade não é observável. Apenas o resultado da decisão - aderir ou não ao tratado - é observável, de forma que supomos que  $y_{i,t} = 1$ , se  $z_{i,t} > 0$ , e  $y_{i,t} = 0$  caso contrário.

A distribuição logística é tal que:

$$P(\varepsilon_{i,t} < x) = \text{logit}^{-1}(x), \text{ para todo } x.$$

Portanto, é fácil derivar conhecida equação da regressão logística.

Outra formulação equivalente é assumir que a variável dependente observada,  $y_{i,t}$  segue uma distribuição de Bernoulli com probabilidade de sucesso  $p_{i,t}$  e que  $\text{logit}^{-1}(p_{i,t}) = X_{i,t} \beta$ . Nessa última formulação, as suposições mais críticas para o nosso trabalho são explicitadas:

---

<sup>12</sup> Para simplificar a exposição, nos concentraremos aqui no modelo logit. Porém, a extensão para o modelo probit é trivial e os resultados são similares

i) as decisões de um ano para outro e entre países são independentes e identicamente distribuídas; ii) O sucesso (isto é, aderir ao tratado) em período anterior não altera a probabilidade de sucesso do período seguinte.

Nessa formulação, parece claro que o modelo logit não é o mais ideal para o nosso tipo de dado. Em primeiro lugar, a depender de onde começamos a contar a data em que um tratado está disponível para adesão, os dados podem ser coletados de tal modo que inicialmente todos os países não aderem no ano 1. Uma forma de evitar isso é colocar o ano inicial da base como o primeiro no qual o tratado está aberto à adesão. Contudo, nós estudamos apenas tratados concluídos, isto é, estudamos a adesão a tratados criados, e desconsideramos todos os possíveis tratados cujas negociações nunca foram concluídas. Assim, uma vez terminadas as negociações do tratado, é provável que muitos países adiram no primeiro momento por terem participado de negociações prévias e como condição de conclusão do tratado.

Problema mais sério, contudo, está em que, embora possa ser uma aproximação razoável assumir que a decisão de um país não aderir a um tratado numa dado ano é independente da decisão de aderir ou não no período seguinte, condicional aos valores dos preditores, essa suposição é claramente inadequada em se tratado da decisão de um país de continuar ou não num tratado, após ter aderido a ele num período anterior. De fato, países raramente denunciam um tratado.

Um modelo mais realista para esse processo de geração dos dados deve considerar que a denúncia de um tratado tem um custo alto, porém não observável. Assim, seria necessário incluir o custo de denúncia do tratado no modelo. Porém, essa variável não é observável e deve ser considerada uma variável latente. Um modelo com efeitos específicos por país poderia estimar essa variável não observável, desde que esse custo não vá no tempo, mas apenas por país.

Não obstante, esse modelo tem de incluir como variável preditora uma variável binária indicando se o país aderiu no período anterior. A inclusão de variável dependente binária defasada como preditora modifica a interpretação dos resultados, já que se passa a estimar um modelo dinâmico. Além disso, induz violação da hipótese de exogeneidade estrita (Morton, 2010).

Mas o principal problema para a nossa abordagem é que não há variação no comportamento de um país após a adesão. Isso significa que uma variável dependente defasada é quase que

perfeitamente correlacionada para todos os países. Isso pode induzir problemas de separação quase perfeita (ver Gelman e Hill, 2007) e explicará quase toda a variância dos dados.

Assim, decidimos adotar outra estratégia de estimação do modelo. Seguindo a abordagem de Bernauer, Kalbhen, Koubi e Spilker (2010), os países só permanecem na base de dados até o momento em que aderem a um tratado, se o fazem. Após a data de adesão, o país é retirado da amostra, de forma que para cada país temos no máximo uma única observação de adesão. Assim, por exemplo, se o Brasil adere a um tratado em 1992, então teremos observações do Brasil desde o primeiro período da amostra até 1992. Anos posteriores a essa data não serão incluídos para o Brasil na estimação do modelo. Assim, é possível estimar um modelo com efeitos específicos, que julgamos mais adequado, como argumentaremos em seguida. Além disso, essa estratégia permite identificar mais adequadamente a associação das covariáveis com a probabilidade de adesão.

Na próxima seção discutimos a vantagem de adotar um modelo com efeitos específicos, mais especificamente o modelo hierárquico Bayesiano. Para tanto, fazemos uma apresentação sobre a inferência Bayesiana antes de entrarmos propriamente nos modelos hierárquicos.

## **4.2. Inferência Bayesiana**

A Inferência Bayesiana tem sido objeto de crescente aplicação nas ciências sociais, incluindo a ciência política (Grimmer, 2010, Pang, 2010, Jackman, 2009, 2004, Shor, Keele, Park, 2007, Bafumi, Gill, 2008, Gill e Walker, 2005). A título de ilustração dessa tendência recente, a edição da revista *Political Analysis* de outubro de 2010 trazia dois artigos com a abordagem Bayesiana já expressa no título. Porém, pouco se sabe sobre a inferência Bayesiana em Ciência Política no Brasil. Além disso, a inferência Bayesiana utiliza alguns conceitos que, ainda que parecidos com os conceitos clássicos, são diferentes.

Incerteza na análise frequentista é baseada em geral no erro padrão, intervalo de confiança e p-valor. Na análise Bayesiana, os conceitos mais utilizados são a probabilidade *a posteriori* e os intervalos de credibilidade. Assim, é importante apresentar esses conceitos principais e a perspectiva Bayesiana sobre inferência.

O restante da seção está organizado como segue: apresentamos inicialmente uma breve história da estatística Bayesiana, os fundamentos filosóficos e pressupostos do Bayesianismo e, por fim, descrição de como se faz a estimação em modelos simples e em modelos hierárquicos.

A disciplina da estatística é bastante moderna e se iniciou mais ou menos na mesma época da maior parte das ciências sociais. Os trabalhos de Laplace (1774, 1781) e Gauss (1809, 1823) são exemplos do começo da estatística. Mas foi no final do século XIX e começo do século XX que as modernas técnicas de inferência estatística foram desenvolvidas. A síntese Neyman-Pierson-Fisher, ainda que com algumas contradições, é a forma mais empregada dentro das ciências sociais quando se trata de inferência estatística. No assim chamado quadro clássico, o objetivo da inferência estatística é “determinar que generalizações sobre a população podem fazer-se a partir da amostra que da mesma foi recolhida” Paulino(2003). Nessa visão, é crucial a distinção entre amostra e população, e a idéia de que cada observação da amostra é coletada sob condições similares.

Como, porém há uma variabilidade de amostra para amostra, a generalização depende não apenas dos dados observados, mas também dos possíveis conjuntos de dados que poderiam ter sido coletados (mas não foram). Daí que os dados amostrais são considerados como uma variável aleatória governada por alguma distribuição de probabilidade. Uma vez que o pesquisador propõe um modelo (isto é, uma distribuição de probabilidade) para a amostra, em geral assume-se que as  $n$  variáveis aleatórias  $X_1, X_2, \dots, X_n$  são independentes e identicamente distribuídas (i.i.d.) condicionalmente a um parâmetro  $\theta$  (ou vetor de parâmetros). A densidade pode então ser representada por:

$$f(x|\theta) = \prod_{i=1}^n f(x_i|\theta)$$

A partir desse modelo, definido pelo pesquisador como o mais apropriado para o processo de geração dos dados, procede-se propriamente à inferência. Mais especificamente, os frequentistas perguntam-se, assumindo que a especificação está correta, o que podemos inferir sobre o verdadeiro parâmetro. Em geral, utiliza-se uma função da amostra, o estimador, como expressão da inferência (pontual) sobre o verdadeiro valor do parâmetro. Por exemplo, seja  $T(X_1, X_2, \dots, X_n)$  tal estimador. Os frequentistas irão apontar  $T$  como estimativa pontual de  $\theta$  e estão interessados nas propriedades desse estimador. Dentre as propriedades mais importantes estão a de o estimador ser não-viesado e consistente. Por exemplo, é comum a preocupação com estimadores viesados de parâmetros.

Para o paradigma clássico ou frequentista, é crucial o princípio da amostragem repetida. Daí que as propriedades dos estimadores são quase sempre avaliadas assintoticamente, isto é, para

um número infinito de repetições da amostra. E a noção de probabilidade será definida a partir do limite da frequência relativa.

O paradigma Bayesiano, ao contrário, é fundado sobre outras bases. O trabalho de Thomas Bayes, *An essay Towards Solving a Problem in the Doctrine of Chances*, publicado postumamente em 1763 por Richard Price, estabeleceu o que é hoje conhecido como Teorema de Bayes.

$$f(\theta|y) = \frac{f(y|\theta)*f(\theta)}{\int_{\Theta} f(y|\theta)*f(\theta)}$$

Em que  $f$  denota apenas a função densidade de probabilidade;  $y$  é (uma) observação, e  $\theta$  é o parâmetro a ser estimado. Um exemplo simples, mas clássico, para ilustrar o teorema de Bayes é dado pelos testes médicos. Nesse caso, substituímos o parâmetro teta ( $\theta$ ) por um evento aleatório (estar ou não doente). Se eu estou interessado em saber a probabilidade de alguém ter uma doença, dado que o teste dessa pessoa é positivo para a doença, então essa probabilidade é dada por:

$$Pr(doente|teste\ positivo) = \frac{Pr(teste\ positivo|doente)*Pr(doente)}{Pr(teste\ positivo|doente)*Pr(doente)+Pr(teste\ positivo|n\~{a}o-doente)*Pr(n\~{a}o-doente)}$$

O denominador pode ser interpretado como a probabilidade de alguém em geral testar positivo, já que a soma das probabilidades de um teste positivo em doentes (verdadeiro positivo) e um teste positivo em não-doentes (falso positivo) é o mesmo que a probabilidade de um positivo. No numerador, a probabilidade de alguém estar doente é designada por probabilidade *a priori*. É assim chamada porque esta é a probabilidade de interesse original, porém reflete a probabilidade antes de obter alguma informação (no caso se o teste é positivo ou não). Como é uma probabilidade antes de obter informação ou dado, consideramos que é uma probabilidade *a priori*. A probabilidade de um teste positivo, dado que a pessoa é doente, é denominada de verossimilhança ou modelo para as informações ou dados a serem obtidos. A verossimilhança, portando, descreve as probabilidades de obter a informação obtida ou observar os dados coletados, dado o parâmetro de interesse (nesse caso, um evento de interesse, estar doente). A probabilidade de alguém estar doente dado que o teste é positivo,  $Pr(doente|teste\ positivo)$ , é denominada probabilidade *a posteriori*. Como a probabilidade de estar doente, sem informação alguma, é *a priori*, após a obtenção da

informação ou dado de que a pessoa testou positiva, a probabilidade de estar doente após essa informação é *a posteriori*, ou seja, após coletar os dados.

Este teorema, um resultado relativamente simples derivado a partir de probabilidades condicionais, não tem sua validade matemática objeto de questionamento, nem tampouco é objeto de controvérsia. A axiomatização matemática da probabilidade tal como exposta classicamente por Kolmogorov é perfeitamente compatível com o teorema de Bayes. Como sói acontecer em disciplinas aplicadas, a controvérsia está na interpretação de partes do teorema e na sua utilização na inferência estatística. Essa situação, diga-se de passagem, não é de todo surpreendente, na medida em que a inferência é uma tentativa de enfrentar de forma rigorosa e sistemática o problema da indução: como eu posso inferir sobre o futuro apenas a partir de observações sobre o passado?

Na interpretação Bayesiana, a densidade *a priori*  $f(\theta)$  reflete a incerteza subjetiva do indivíduo antes de coletar os dados. Dessa forma, a probabilidade é interpretada como uma medida dos graus de crença de um indivíduo sobre um determinado evento ou parâmetro (ou vetor de parâmetros). Pode-se pensar assim que o teorema de Bayes especifica matematicamente como se deve proceder à atualização de opiniões e/ou crenças após obter informações sobre o objeto de interesse.

Para os nossos propósitos, importa salientar que a estimação Bayesiana se dá, portanto, por meio de uma especificação de uma distribuição de probabilidade dita *a priori*, a especificação de uma verossimilhança ou – no caso presente, um modelo de regressão - e a obtenção de uma densidade *a posteriori* para os parâmetros de interesse. *A posteriori* representa a incerteza sobre a distribuição do parâmetro. Assim, ao invés de termos de olhar a distribuição amostral para caracterizar a incerteza como fazemos quando computamos o p-valor de uma estatística, na inferência Bayesiana a própria *posteriori* quantifica nossa incerteza.

Assim, a densidade *a posteriori* dos parâmetros representa toda a incerteza envolvida na estimação do parâmetro. Medidas resumo, como a média, são estimativas pontuais da densidade e permitem resumir nossa incerteza. Do mesmo modo, podemos computar intervalos de credibilidade, que nada mais são do que intervalos definidos da densidade. Assim, um intervalo de credibilidade de 95% de certeza nada mais é do que quais pares de valores dos percentis 2,5% e 97,5% da densidade *a posteriori*. Em outras palavras, diferentemente do intervalo de confiança, que é complicado de ser entendido, um intervalo de

credibilidade de 95% nos diz que o parâmetro está naquele intervalo com 95% de probabilidade.

Num contexto de regressão simples,  $y_i = a + b * x_i + \varepsilon_i$ ,  $\varepsilon_i \sim N(0, \sigma^2)$ , a obtenção da *posteriori* se dá do seguinte modo: reescrevemos a equação de regressão:

$$(y_i | a, b, x_i, \sigma^2) \sim N(a + b * x_i, \sigma^2),$$

Essa é nossa verossimilhança, isto é, o modelo probabilístico para o processo de geração dos dados. E nós estamos interessados em obter uma distribuição de probabilidade para os parâmetros  $a$  e  $b$ , e eventualmente  $\sigma^2$  se este não for conhecido (em geral não é). Devemos portanto especificar uma densidade *a priori* para os parâmetros e em seguida aplicar o teorema de Bayes. Assumindo para facilitar que a variância é conhecida, e supondo que não tenho informação *a priori* sobre os possíveis valores de  $a$  e  $b$ , exceto que podem ser positivos ou negativos e é improvável que sejam maiores do que 10, podemos representar essa nossa opinião *a priori* sobre os parâmetros por:

$$a \sim N(0, 1000)$$

$$b \sim N(0, 1000)$$

A obtenção da *posteriori* para os parâmetros, mesmo nesse caso simples, é matematicamente complicada e não derivaremos aqui. Importante é que, uma vez obtida a *posteriori* para os parâmetros  $a$  e  $b$ , pode ser útil reportar o valor médio da *posteriori* de  $a$  e  $b$  e um intervalo de credibilidade de 95% para essas estimativas. Suponhamos que a média da *posteriori* de  $a$  é 2,3 e com 95% de probabilidade o parâmetro  $a$  está entre 2 e 2,6, então este é o intervalo de credibilidade a 95% por cento para o parâmetro  $a$ . Não somente a interpretação é similar à prática tradicional, como ela é direta e não requer referências a repetições de eventos em condições similares associadas ao intervalo de confiança. No Bayesianismo, o intervalo de credibilidade significa a probabilidade de o parâmetro estar contido dentro do intervalo especificado.

Além disso, como o intervalo de credibilidade é construído a partir da densidade *a posteriori*, os intervalos não precisam ser simétricos. Se a densidade não for simétrica, os intervalos também não serão simétricos. Por fim, é possível calcular diretamente a probabilidade de o valor do parâmetro ser positivo ou negativo.

Uma preocupação natural na inferência Bayesiana é o papel da priori. Como a posteriori é uma combinação da priori com a verossimilhança, há críticas ao fato da inferência Bayesiana permitir que a opinião subjetiva do pesquisador influencie nas conclusões. Todavia, quanto mais dados forem coletados e incluídos na análise, menor a importância da priori na forma final da posteriori. Mas para discutir mais adequadamente o papel da priori, é importante ter em mente o teorema da representação, devido a de Finetti.

Naturalmente, é possível que o leitor se pergunte em que sentido podemos falar de parâmetros no Bayesianismo. No paradigma clássico, parâmetros são valores populacionais, fixos, mas desconhecidos. Para os Bayesianos, parâmetros são variáveis aleatórias caracterizadas por uma distribuição de probabilidade. Se porém a probabilidade é subjetiva e pode variar de indivíduo para indivíduo, em que sentido podemos falar de parâmetros numa análise Bayesiana? Como é possível então utilizar o conceito de parâmetro no paradigma Bayesiano? O teorema da representação, de Bruno de Finetti, é uma forma de justificar essa utilização do termo parâmetro. Devido à importância para os modelos hierárquicos da idéia de permutabilidade, associada ao Teorema da Representação, apresentamos a seguir, de forma resumida, os resultados de de Finetti.

A noção de permutabilidade, central na análise Bayesiana, é parecida com a idéia de variáveis aleatórias independentes e identicamente distribuídas, porém mais geral. De Finetti irá introduzir a noção de permutabilidade, de forma a preservar a simetria dos eventos, e ao mesmo tempo captar alguma relação de dependência entre as observações e podermos aprender com a experiência. Seguindo (Schervish, 1995), a definição de permutabilidade pode ser dada como segue:

Um conjunto finito  $X_1, \dots, X_n$  de quantidade aleatórias é dita permutável se cada permutação  $(X_1, \dots, X_n)$  tem a mesma distribuição conjunta de toda outra permutação. Uma coleção infinita é permutável se cada subcoleção finita é permutável.

Na prática, a noção de permutabilidade é parecida com aquela de i.i.d. condicional a algo (parâmetro), e reflete a idéia de que eventos são simétricos, sem precisar entrar na seara de independência ou de uma frequência relativa de frequências. De fato, é possível estabelecer formalmente a relação entre i.i.d. condicional a algo (parâmetro) e permutabilidade. Como não é nosso interesse discutir a fundo essas questões teóricas e de fundamento, iremos aqui apenas apresentar esses resultados. O leitor interessado pode consultar as referências apresentadas.

Tendo apresentado a noção de permutabilidade, passemos ao Teorema da Representação, de De Finetti.

Teorema da Representação de de Finetti para variáveis aleatórias Bernoulli.

Seja  $T_t (t = 1, 2, \dots)$  uma sequência infinita de variáveis aleatórias de Bernoulli indicando sucesso 1 ou fracasso 0 de algum evento de interesse. Para qualquer sequência finita  $Y_t (t = 1, 2, \dots)$  defina o número médio de ocorrências por:

$$\bar{y}_T = \sum_{t=1}^T T_t / T$$

Seja  $h(y_1, y_2, \dots, y_T)$  a probabilidade de massa, e  $H(y)$  a acumulada, então  $h(\cdot)$  tem representação:

$$h(y_1, y_2, \dots, y_T) = \int_0^1 \prod_{t=1}^T \theta^{y_t} (1 - \theta^{1-y_t}) dF\theta$$

$$F(\theta) = \lim_{T \rightarrow \infty} P_H(\bar{y}_T \leq \theta)$$

A prova do teorema pode ser encontrada em Schervish (1995). Concentremo-nos na interpretação do teorema. Ele diz que, condicional em  $\theta$ ,  $h$  é uma i.i.d. Bernoulli com parâmetro  $\theta$ . Nos dizeres de Koope et. al. (2007, p. ??):

*The theorem (...) implies that it is as if, given  $\theta$ ,  $Y_t (t = 1, 2, \dots, T)$  are iid Bernoulli trials where the probability of a success is  $\theta$ , and the 'parameter'  $\theta$  is assigned a probability distribution with c.d.f  $F(\cdot)$  that can be interpreted as a belief about the long-run relative frequency of  $\bar{y}_T \leq \theta$  as  $T \rightarrow \infty$ . From de Finetti's standpoint, both the quantity  $\theta$  and the notion of independence are 'mathematical fictions' implicit in the researcher's subjective assessment of arbitrary long observable sequences of successes and failures. The parameter  $\theta$  is of interest primary because it constitutes a limiting form of predictive inference about the observable  $\bar{y}_T$ . The mathematical construct  $\theta$  may nonetheless be useful. However, the theorem implies that the subjective probability distribution need not apply to the 'fictitious  $\theta$ ' but only to the observable exchangeable sequences of successes and failures.*

Como se observa pela longa citação, tanto o parâmetro  $\theta$  quanto a noção de independência são apenas construtos matemáticos. Nós podemos pensar apenas nos dados observáveis ou então em parâmetros, pois matematicamente são equivalentes (no sentido acima). Em outras palavras, o teorema da representação nos habilita a trabalhar com a idéia de parâmetros e reter a noção de probabilidade subjetiva. Ademais, como diz José bernardo (1994, p. ??):

*if a sequence of observations is judged to be exchangeable, then, any finite subset of them is a random sample of some model  $p(x_i|\theta)$ , and there exist a prior distribution  $p(\theta)$ , which has to describe the initially available information about the parameter which labels the model.*

Em última instância, então, a inferência Bayesiana consiste apenas em estabelecer uma priori para o parâmetro de interesse, obter dados, formular a verossimilhança dos dados, isto é, modelar a dependência entre os dados condicional ao parâmetro de interesse e, por meio do teorema de Bayes, obter a densidade a posteriori para o parâmetro.

Como o objetivo da inferência bayesiana é calcular probabilidades posteriores a partir dos dados, tudo que se tem que fazer é modelar explicitamente as relações entre as variáveis e fornecer priores para os parâmetros. Feito isso, basta calcular as probabilidades a posteriori e reportar sumários da densidade a posteriori encontrada.

É verdade que essa simplicidade dependia, em tempos passados, da capacidade do pesquisador de realizar operações de integração complicadas e difíceis de serem resolvidas analiticamente. Porém, com o advento dos modernos computadores com alta capacidade computacional, é possível resolver as integrais por meio de simulações, usando principalmente o assim chamado método de MCMC - *Markov Chain Monte Carlo* (Jackman, 2000).

MCMC, como a própria sigla aponta, é composta de dois “MCs”, Cadeias de Markov e Monte Carlo<sup>13</sup>. O princípio de Monte Carlo afirma que “anything we want to know about a random variable  $\theta$  can be learned by sampling many times from  $f(\theta)$ , the density of  $\theta$ ” (Jackman, 2009, p. 134). Em outras palavras, de posse da distribuição de probabilidade para  $\theta$ , podemos calcular a média, moda, mediana, desvio-padrão, percentis etc. por meio da amostragem simulada de  $n$  observações **independentes** a partir da distribuição de  $\theta$ .

Assim, por exemplo, suponha que  $\theta \sim \text{logística}(1, 1)$ . É possível derivar, analiticamente, a média de  $\theta$ ,  $E[\theta]$ , e a variância de  $\theta$ ,  $Var[\theta]$ . Porém, se tivermos uma forma de simular valores de acordo com essa distribuição para  $\theta$ , então podemos calcular simplesmente  $E[\theta]$  e  $Var[\theta]$  do seguinte modo: Faça  $n$  simulações de  $\theta$  com base em *logística* (0, 1), e compute

---

<sup>13</sup> Descrevemos brevemente aqui em que consistem os dois “MCs”, já que não é nosso objetivo explicar detidamente como empregar o método de MCMC. O leitor interessado pode consultar Jackman (2009, 2000).

a média amostral<sup>14</sup>  $E[\hat{\theta}] = \frac{\sum_{i=1}^n \theta_i}{n}$ . Pela lei dos grandes números, quanto maior  $n$ , melhor a aproximação, isto é,  $\frac{\sum_{i=1}^n \theta_i}{n} \rightarrow E[\theta]$  à medida que  $n \rightarrow \infty$ . O gráfico 4.2 apresenta a média amostral de  $\theta$  para  $n = 10, 100, \dots, 10.000.000$ . Como se pode perceber, à medida que aumentamos o número de simulações ou iterações, o valor converge para 1, que é o valor verdadeiro. As limitações para a aproximação são, apenas, de ordem computacional. Quanto maior o número de simulações, mais tempo o computador leva. Além disso, os números reais são infinitos e o computador não consegue distinguir todos os números, de forma que a aproximação nunca poderá ser perfeita. De todo modo, para efeitos práticos, a aproximação pode ser tão boa quanto se queira e, de fato, funciona na prática.

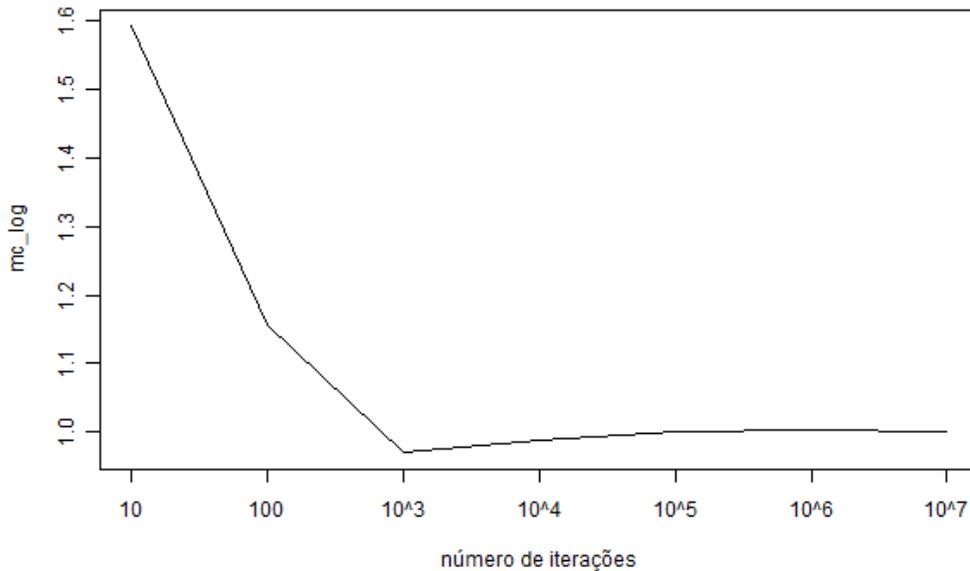


Gráfico 4.2. Média amostral de  $\theta$ , estimada por simulação de Monte Carlo para  $n = 10, 100, \dots, 10.000.000$ . O valor verdadeiro de  $\theta$ , derivado analiticamente, é 1. Com 10 simulações, a aproximação é ruim. A partir de 1000 simulações, a aproximação é bastante boa e converge para o valor verdadeiro.

Não obstante, apesar da eficácia das simulações de Monte Carlo, ela depende crucialmente de sabermos qual a distribuição de probabilidade alvo ou de interesse e de sermos capazes de realizar simulações no computador a partir de uma dada densidade. Porém, na maior parte dos casos, não temos como realizar simulações diretamente da densidade de interesse, em geral porque a função densidade não é de nenhuma das famílias de distribuições conhecidas

<sup>14</sup> A rigor, temos não a média amostral, mas a média populacional (Casela e George, 1992). Mas essa distinção não é muito importante aqui.

(como a Normal, Uniforme, Cauchy, Dirilichet, Pareto, Poisson, Binomial entre outras). Assim, o princípio de Monte Carlo não pode ser aplicado para os casos mais gerais e mais interessantes.

Na análise Bayesiana, em geral estamos interessados na distribuição *a posteriori* conjunta de vários parâmetros (intercepto, inclinação, variância etc.),  $p(\boldsymbol{\theta}|\text{dados})$ . Como temos vários parâmetros, a distribuição é multidimensional e aplicar o princípio de Monte Carlo diretamente é em geral inviável. Porém, pode ser possível simular valores a partir das distribuições a posteriori condicionais, posto que é um fato estabelecido que a densidade conjunta é completamente caracterizada pelas densidades condicionais e marginais.

Considere o caso simples de duas variáveis aleatórias,  $X$  e  $Y$ . Partindo de um valor inicial arbitrário de  $Y$ ,  $y_0$  (mas dentro dos valores possíveis de  $Y$ ), condicional ao valor de  $Y$ , podemos simular a partir de  $\Pr(X|Y = y_0)$  e obter um valor  $x_1$  de  $X$ <sup>15</sup>. Uma vez obtido  $x_1$ , podemos então simular uma realização de  $Y$ , condicional ao valor de  $X$  obtido anteriormente,  $\Pr(Y|X = x_1)$ . Tendo um novo valor  $y_1$ . Com esse novo valor, podemos obter um novo valor para  $X$ , e assim por diante, quantas vezes quisermos. Esse procedimento, denominado de Amostrador de Gibbs, permite simular valores de  $X$  e  $Y$ .

Contudo, os valores amostrados não são independentes, pois na iteração  $T$  eles dependem do valor amostrado em  $T - 1$ . Não obstante, é possível mostrar que sob certas condições esses valores simulados constituem uma Cadeia de Markov com a propriedade de que convergem para uma distribuição estacionária. Uma Cadeia de Markov que converge para uma distribuição estacionária tem a característica de que, não importa os valores iniciais da Cadeia, na convergência a distribuição de probabilidade é invariante e valores simulados a partir dessa distribuição são, portanto, independentes.

Assim, temos que valores amostrados pelo Amostrador de Gibbs, após convergirem para a distribuição estacionária da Cadeia de Markov, constituem amostras independentes das variáveis aleatórias, de forma que podemos aplicar o princípio de Monte Carlo após a convergência. Temos então uma simulação por *Markov Chain Monte Carlo*.

Há é claro outros métodos além do Amostrador de Gibbs que utilizam MCMC. Importa aqui apenas assinalar que, qualquer que seja o método de MCMC, ele utiliza o fato de que após

---

<sup>15</sup> Normalmente, ‘ $X$ ’, em caixa alta, indica a variável aleatória, e ‘ $x$ ’, em caixa baixa, indica uma realização (uma observação) de  $X$ .

atingir a convergência da Cadeia de Markov na distribuição estacionária, podemos aplicar o método de Monte Carlo.

Por isso, na estimação Bayesiana, é preciso checar a convergência da Cadeia para garantir que os valores amostrados sejam efetivamente independentes. Há vários testes para se analisar convergência de Cadeias de Markov, mas não é possível provar por meio dos testes que uma Cadeia efetivamente convergiu, apenas rejeitar a hipótese de que temos evidência de não convergência. Além disso, é preciso indicar um número de iterações para serem descartadas (*burn-in*), antes de a convergência ser atingida. Assim, em estimações Bayesianas por MCMC, temos que garantir que as Cadeias convergem e definir um período de *burn-in*. No presente trabalho, as Cadeias tinham no mínimo 50.000 iterações e máximo de 300.000 iterações, com período de *burn-in* de 50% do total de iterações. A checagem da convergência foi realizada por meio da visualização gráfica do *trace-plot* de três cadeias partindo de pontos iniciais diferentes e por meio da estatística de Gelman-Rubin (Jackman, 2009; Gelman, Carlin, Stern e Rubin, 2004; Brooks e Gelman, 1997; Gelman e Rubin, 1992).

Assim, atualmente a principal dificuldade para a aplicação dessa técnica é o desconhecimento da abordagem Bayesiana e de como utilizar softwares que, por meio de MCMC, estimem a densidade posterior de interesse. No caso de modelos mais complicados do que simples estimação de uma média ou proporção, a modelagem que é requerida do pesquisador é basicamente a mesma para a análise de regressão clássica, tendo como adicional apenas especificação de priores para os parâmetros.

Para nós, o conceito de permutabilidade é particularmente esclarecedor quando se trata dos assim chamados modelos hierárquicos. Os modelos hierárquicos são uma classe especial de modelos bayesianos, em que a estrutura dos dados é levada em consideração (Gelman e Hill, 2007). Ou seja, se o pesquisador acredita que o intercepto ou a inclinação do modelo varia por subgrupos da amostra, então essa hierarquia pode e deve ser incorporada ao modelo. Modelos com dados com características de *cross-section* e série de tempo, isto é, dados em painel, também podem ser modelados com análise multinível. Apresentemos então, em mais detalhe, o que é um modelo Bayesiano Hierárquico ou Multinível.

### **4.3. Modelos Hierárquicos**

Uma das vantagens da modelagem Bayesiana é a possibilidade de naturalmente construir modelos hierárquicos. O objetivo dessa nota é explicar a vantagem dos modelos hierárquicos Bayesianos. É comum, especialmente em ciências sociais, que alguma variável aleatória possa

ser agrupada em conjuntos similares. Por exemplo, ao observarmos as notas dos alunos em várias escolas, é natural considerar que as notas dentro de cada escola estão correlacionadas. Nesse caso, nós gostaríamos de poder modelar essa correlação em nossa inferência estatística.

Formalmente, suponhamos que temos uma variável aleatória  $y_{ij}$  para os indivíduos  $i = \{1, 2, \dots, n\}$  em  $j = 1, 2, \dots, G$  grupos. Se os valores de  $y_{ij}$  dentro de cada grupo  $j$  são correlacionadas, então não podemos assumir simplesmente que temos observações independentes (ou, nos termos Bayesianos, permutáveis). Assumir que as observações são independentes é pressupor que temos mais graus de liberdade do que efetivamente temos, e, ao mesmo tempo, desprezar uma fonte de informação presente nos dados.

Então, não podemos dizer que a sequência  $y_{ij}$  é permutável. Porém, pode ser que, condicional ao grupo a que pertence, os  $y_{ij}$  sejam permutáveis. Isto é, dentro de cada grupo, os valores  $y_{ij}$  são independentes e identicamente distribuídos, condicional ao parâmetro relevante. Isto é, condicional aos vetores de parâmetros de cada grupo, as  $y_{ij}$  são permutáveis. Mas não somente é possível pensar que as sequências são permutáveis dentro de cada grupo, mas também que cada grupo é permutável entre si.

Ou seja, não somente podemos modelar e estimar uma posteriori para os parâmetros que determinam os valores individuais da variável  $y$ , mas também podemos estimar os parâmetros que determinam a variabilidade entre grupos. Retomando o exemplo inicial da seção, podemos estimar não somente a média dos alunos, mas também a média de cada escola, de forma que é possível avaliar se são as características individuais de cada aluno ou as características de cada escola que mais pesam na determinação da notas dos mesmos.

Assim, uma das principais vantagens dos modelos Bayesianos hierárquicos é que as estimativas obtidas para a média de cada indivíduo levam em consideração tanto a informação no interior de cada grupo como também a informação entre grupos. Em outras palavras, tudo se passa como se tivéssemos um estimador de *shrinkage*, que é uma média ponderada dos resultados dos indivíduos  $i$  e a média dos grupos  $j$ . Esse é um resultado bem estabelecido e facilmente verificável para modelos simples como modelos hierárquicos lineares com erro normalmente distribuído (veja Koop, Poirier e Tobias, 2007). Essa propriedade dos modelos hierárquicos Bayesianos é útil também na medida em que ajuda a resolver o problema das múltiplas comparações entre grupos. Ao invés de corrigir as

comparações aumentando o intervalo de confiança, as estimativas são agrupadas em torno da média comum (ver Gelman e Hill, 2007, Gelman, 2009).

No presente trabalho, nosso interesse é estimar um modelo hierárquico para variável dependente binária. Assim, apresentamos o modelo hierárquico a ser estimado.

#### 4.4. Modelo Hierárquico Logístico

No presente trabalho, nós estamos interessados em modelar os determinantes da variável aleatória adesão a tratados de patentes pelo país  $i$  no ano  $t$ . Desse modo, genericamente temos:

$$\Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha + b * x_{it} )$$

Porém, é razoável assumir que a decisão de um país aderir ou não num dado período de tempo é correlacionada com a decisão em períodos anteriores, condicional ao valor das covariáveis. Nesse caso, podemos pensar em cada país como um grupo, de modo que, condicional ao grupo a que pertencem, as observações de cada país são permutáveis no interior dos países. Formalmente, tal modelo hierárquico pode ser escrito do seguinte modo:

$$\Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha_i + b_0 + b * x_{it} )$$

$$\alpha_i \sim \text{Normal} (\mu_{\text{alpha}}, \sigma_{\text{alpha}}^2)$$

No modelo Bayesiano, os  $\mu_{\text{alpha}}$  e  $\sigma_{\text{alpha}}^2$  são também chamados de hyperparâmetros, pois podem ser interpretados do seguinte modo. Inicialmente, sorteiam-se aleatoriamente os efeitos específicos de cada país, a partir da distribuição Normal com parâmetros  $\mu_{\text{alpha}}, \sigma_{\text{alpha}}^2$ . Uma vez sorteados os efeitos específicos,  $y_{it} \sim \text{Bernoulli} (\alpha_i + b_0 + b * x_{it})$ . Ou seja, os hyperparâmetros governam as realizações dos efeitos específicos, que por sua vez governam as realizações de  $y_{it}$ .

Portanto,  $\sigma_{\text{alpha}}^2$  representa a variância populacional entre países. Se esta for muito grande, isso significa que as diferenças observadas na variável  $y_{it}$  são maiores entre países que no interior do próprio país.

Naturalmente, é possível estender o modelo hierárquico acima com a inclusão de mais hierarquias - por exemplo, se os  $\alpha_i$  forem agrupados em supergrupos, como por exemplo,

continentes -, então teremos uma densidade para esses supergrupos com os respectivos (hyper)hyperparâmetros. Além disso, uma vez que todos os parâmetros são modelados, é possível incluir preditores dos hyperparâmetros. Assim, por exemplo, o modelo acima poderia ser escrito da seguinte forma:

$$\Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha_i + b_0 + b * x_{it} )$$

$$\alpha_i \sim \text{Normal} (g_0 + g_1 * w_i, \sigma_{\alpha}^2),$$

em que  $w_i$  é um preditor que varia apenas entre países, mas não intra-país. Tal preditor pode ser o tipo de colonização do país, data de independência, entre outros.

Outra possibilidade é assumir que, num dado ano, as decisões dos países de aderirem ou não ao tratado são correlacionadas entre si. Ou seja, os efeitos específicos variam por ano, mas não entre países. Em outras palavras, os países estão sujeitos aos mesmos choques aleatórios no tempo. Porém, condicional ao choque, são os preditores de cada país que determinam a decisão deles aderirem ou não ao tratado. No caso de tratado de patentes, essa é uma hipótese bastante razoável, já que muitos choques temporais ocorreram no período e que influenciaram as decisões dos países. Nos anos 80, as pressões dos interesses organizados em todo o mundo para fortalecer a proteção de patentes podem ser consideradas generalizadas. Nos anos 90, a criação da OMC e o tratado TRIPs, que versa também sobre patentes, exerceram uma força poderosa sobre os países. Do mesmo modo, o crescimento da interdependência no tempo aumentou, supostamente, os custos dos países não participarem do regime de propriedade intelectual. Por fim, o fim da Guerra Fria alterou a percepção dos países sobre as vantagens e desvantagens de aderir aos regimes internacionais. Formalmente, um modelo com efeitos específicos por tempo poderia ser escrito do seguinte modo:

$$\Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha_t + b_0 + b * x_{it} )$$

$$\alpha_t \sim \text{Normal} (\mu_{\alpha}, \sigma_{\alpha}^2)$$

Outra forma de pensar um modelo hierárquico é como um caso intermediário entre uma regressão *pooled* e *unpooled*. Na regressão *unpooled*, estimam-se os efeitos específicos rodando uma regressão separada para cada grupo. Na regressão *pooled*, estimamos uma regressão para todos os dados, com o mesmo intercepto para todos os grupos. No primeiro caso, estamos assumindo que cada grupo é diferente entre si e nenhuma informação pode ser obtida dos demais grupos para estimação dos resultados. No segundo caso, estamos

assumindo que os grupos são todos iguais e nenhuma informação pode ser obtida pela comparação entre grupos. O caso intermediário é justamente o modelo hierárquico. E a variância entre grupos,  $\sigma_{\alpha}^2$ , estimada pelos dados, indicará se os dados estão mais próximos do modelo *pooled* ou *unpooled*. No limite, se  $\sigma_{\alpha}^2 \rightarrow \infty$ , os grupos são totalmente diferentes e temos o caso *unpooled*. Se  $\sigma_{\alpha}^2 \rightarrow 0$ , não há variação entre grupos e o modelo *pooled* é mais adequado.

No presente trabalho, optamos por estimar um modelo hierárquico com efeitos específicos por tempo, e não por país. A razão pela qual não estimamos efeitos específicos por país é que devido à baixa variabilidade dos dados, países que nunca aderiram ao tratado teriam seu comportamento explicado fortemente pelo efeito específico. Apesar de o modelo hierárquico forçar as estimativas para a média geral (efeito *shrinkage* ou *partial pooling*), ainda assim os valores dos efeitos específicos seriam muito grandes, tornando o efeito dos demais preditores insignificantes, se não estatisticamente, em termos de contribuição para explicação da variância dos dados.

Algumas vezes os modelos hierárquicos são chamados de modelos mistos, pois incluem um efeito fixo  $b_0$  e um efeito aleatório  $a_1$ . Porém, a nomenclatura de efeitos fixos e efeitos aleatórios são também utilizados com outro significado, principalmente na literatura econométrica. Nesse outro significado, efeitos fixos são interceptos específicos, porém, sem modelagem probabilística. Assim, efeitos fixos por países são aquelas características constantes dos países, porém não observadas. E os efeitos aleatórios seriam efeitos específicos por país com modelagem probabilística. Como na estimativa Bayesiana todos os parâmetros são variáveis aleatórias e devem ser modelados, essa definição torna tudo muito confuso. Portanto, preferimos simplesmente nos referir a efeitos específicos, sem entrar na denominação fixo ou aleatório. De fato, Gelman e Hill (2007) evitam o uso do termo efeitos fixos e/ou efeitos aleatórios por considerar confusas tanto a definição dos termos quanto as recomendações sobre quando usar efeitos fixos ou efeitos aleatórios.

## 5. Resultados Empíricos - Convenção de Paris

O objetivo deste capítulo é estimar os determinantes da adesão à Convenção de Paris, artigos 1-12, por meio de um modelo hierárquico Bayesiano. Nós apresentamos evidências de que os principais determinantes da adesão são os pedidos de patentes no exterior e a legislação doméstica. Esses achados corroboram a teoria institucionalista segundo a qual a decisão de adesão está associada aos custos e benefícios específicos dos tratados para os países.

Estados com maior número de patentes no exterior tendem a se beneficiar da Convenção de Paris. E países com legislação doméstica mais forte em patentes têm menor custo de *compliance* e necessidade de modificar a legislação doméstica para internalizar as normas da Convenção. Ademais, não achamos evidências de que o regime político doméstico seja relevante e nem evidências fortes sobre a importância dos benefícios colaterais na decisão de adesão. Um resultado contra-intuitivo de nossas estimativas está na associação negativa entre nível educacional de um estado e a probabilidade de adesão. Porém, nós recomendamos cautela nesses resultados, devido a possíveis problemas de viés de seleção, que discutimos mais abaixo.

A discussão dos resultados está organizada como segue: inicialmente apresentamos estatísticas descritivas para os preditores e possíveis problemas de mensuração; em seguida mostramos os resultados da regressão ao nível dos países e os resultados ao nível dos quinquênios. Por fim discutimos o ajuste do modelo.

### 5.1. Métodos e procedimentos empíricos

Com vistas a estimar os determinantes da adesão à Convenção de Paris, tomamos como variável dependente a data de entrada em vigor, por país-ano, dos artigos 1 a 12 da Convenção de Paris, revisada no ato de Estocolmo em 1968. O período analisado compreende os anos de 1970 até 2000, já que a Convenção, tal como revisada em Estocolmo se inicia em 1968 e 2000 é o último ano que temos informações para alguns dos principais preditores como educação e legislação doméstica. Foram utilizados dados quinquenais, de forma que temos dados de adesão em 1970, 1975, 1980, 1985, 1990, 1995 e 2000. No caso de países que aderiram entre quinquênios, é considerada a adesão apenas no próximo quinquênio. Ou seja, se um país aderiu em 1981 ao tratado em questão, então consideremos que o país aderiu ao tratado apenas em 1985. A amostra consiste de 309 observações no formato país-ano, com 89 países desenvolvidos e em desenvolvimento incluídos na análise.

Como queremos estimar os determinantes das decisões de adesão dos estados, temos de enfrentar o problema de que estados raramente denunciam tratados internacionais. A decisão de aderir a um tratado é, portanto, diferente da decisão de permanecer no tratado, tendo já aderido a ele previamente. Devido a essa assimetria na tomada de decisão, retiramos da amostra os países após a adesão, para evitar que contássemos mais de uma vez a decisão de aderir. Conforme discutido no capítulo anterior, de metodologia, essa nos parece a melhor estratégia de estimação. Além disso, está em acordo com alguns estudos similares, que adotaram o mesmo procedimento (Bernauer, Kalbhen, Koubi e Spilker, 2010).

A decisão de utilizar dados quinquenais, e não anuais, deve-se à limitação de informação das variáveis preditoras. As variáveis “nível educacional” e “legislação doméstica de patentes” estavam disponíveis apenas a cada cinco anos. Além disso, a variável “renda per capita”, por conter interpolações, é mais bem adaptada para intervalos de tempo maiores do que análises anuais.

As variáveis preditoras ao nível do país incluídas no modelo são: i) renda per capita; ii) regime político doméstico; iii) nível educacional do país; iv) patentes nos EUA; v) grau de proteção da legislação doméstica de patentes; vi) fluxo de investimento externo direto como proporção do PIB.

Uma preocupação estatística importante é saber se o procedimento de retirar um país da amostra induz a viés nas estimativas. Se uma variável qualquer  $x$  está associada positivamente à probabilidade de adesão, então países com maiores valores de  $x$  nos anos iniciais da amostra tenderão a aderir ao tratado cedo e serão retirados da amostra. Com o passar do tempo, restarão na amostra apenas os países com valores de  $x$  menores. Se, porém, variáveis sistêmicas que influenciam todos os países da amostra criarem uma tendência a que os países participem do tratado, então teremos vários países aderindo ao tratado com valores menores de  $x$ . Então, empiricamente, observaremos uma relação negativa entre  $x$  e a probabilidade de adesão, quando na verdade a associação é positiva.

Para minimizar esse problema, nós iremos comparar para cada variável preditora a distribuição dos dados na amostra com retirada de países após adesão e na amostra completa, sem retirada de países. O objetivo da comparação é tentar apontar evidências qualitativas sobre a provável direção do viés<sup>16</sup>. Embora não possamos quantificar o tamanho desse viés, é

---

<sup>16</sup> Por evidências qualitativas queremos dizer que não realizaremos nenhum teste estatístico

suficiente apontar a direção do viés para qualificar a robustez dos achados no espírito advogado por Manski (2009). Manski (2009) introduz a noção de parâmetros parcialmente identificáveis, segundo a qual ainda que não seja possível estimar pontualmente um parâmetro, pode ser possível estabelecer limites não-triviais sobre o parâmetro. Assim, os intervalos representando incertezas de um parâmetro subestimado apresentam limites inferiores não triviais sobre os parâmetros, embora não possamos quantificar os limites superiores. No caso de parâmetros superestimados, temos limites superiores não triviais, mas não limites inferiores.

Abaixo temos algumas estatísticas descritivas de cada variável preditora na amostra utilizada na modelagem estatística e na amostra completa, além de detalhes do processo de mensuração das mesmas.

## 5.2. Renda per capita

Para renda per capita, nossa *proxy* para nível de desenvolvimento, utilizamos a variável **rgpdl2** da Pen World Table, versão 6.3, que mede a renda per capita em termos de poder de paridade de compra (PPP) e é mais estável que variáveis de versões anteriores, como **rgdpl** (Heston, Summers e Aten, 2009). Como nossos dados são quinquenais, utilizamos a média de cinco anos tomando por ano base 1971, 1976, 1981, 1986, 1991 e 1996 até o quinquênio. Assim, renda per capita de um país em 1980, na nossa base, é a média da renda per capita do país no período 1976-1980. Nós padronizamos a variável durante o processo de estimação, ao subtrair a média de todos os países em todo o período e dividir pelo desvio padrão. A padronização da variável permitiu que a convergência das simulações se desse de forma mais eficiente, de forma que essa especificação foi empregada em todas as simulações.

Na nossa amostra, a renda per capita média é de US\$ 7058 e a mediana é US\$ 5225 e o desvio-padrão é de US\$ 6018. O gráfico 5.2 apresenta um box-plot da renda para os países que aderiram e não aderiram à Convenção de Paris. O Box-plot mostra a dispersão dos dados. O ponto no interior das caixas representa a mediana. As caixas representam o primeiro e terceiro quartil, isto é, 25% dos dados estão abaixo do primeiro quartil e 75% dos dados estão abaixo do terceiro quartil. Pontos fora do intervalo das linhas tracejadas indicam outliers assumindo que a distribuição dos dados fosse Normal.:

---

formal sobre existência, magnitude e direção do possível viés. Iremos apenas utilizar estatísticas descritivas e visualização gráfica para essa análise qualitativa. A análise gráfica qualitativa, porém, não está em oposição à análise quantitativa, mas pode ser visto como um complemento (Gelman 2003a, 2003b).

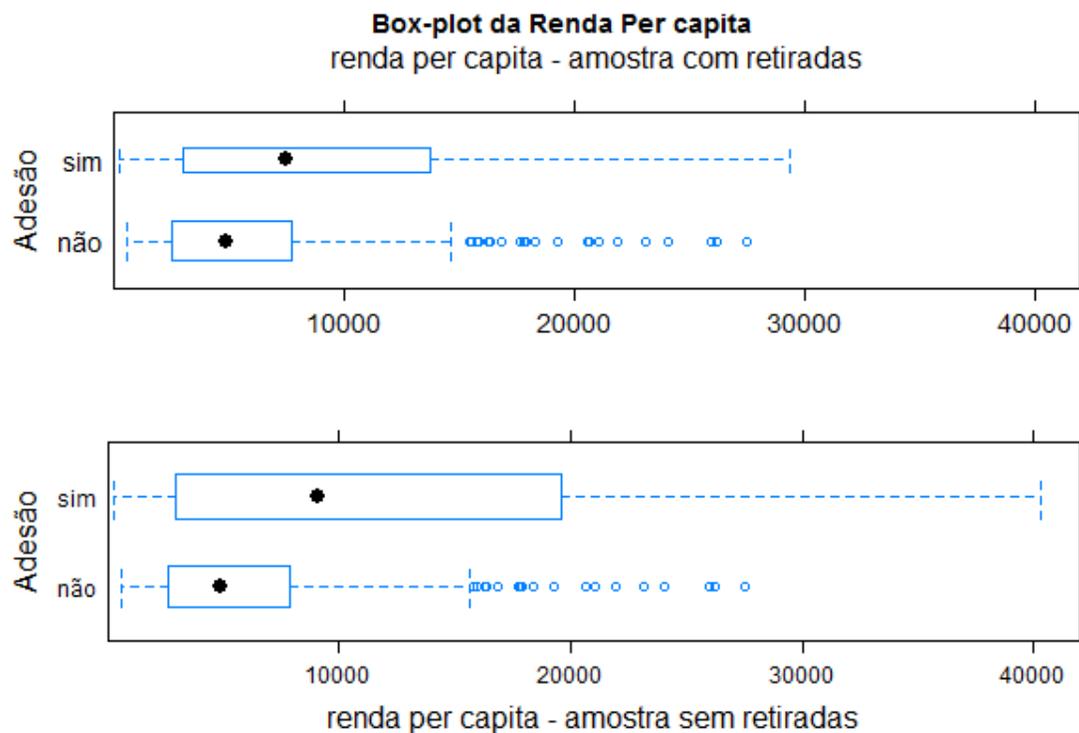


Gráfico 5.2 box plot da renda per capita. O painel superior mostra um Box-plot da renda per capita para os países que aderiram à Convenção de Paris (“sim”, no eixo y), e não aderiram (“não”, no eixo y) na amostra com retiradas, utilizada na regressão. A renda mediana dos países que aderem à Convenção de Paris é maior do que a dos que não aderem. O primeiro quartil de ambos os grupos de países é similar, mas o terceiro quartil da renda dos países que aderem é maior do que entre os não aderentes. Esses dados mostram uma associação positiva entre renda per capita e adesão. A comparação do painel superior (amostra com retiradas) e painel inferior (amostra sem retiradas) mostra que o efeito positivo da renda per capita sobre adesão diminuído na amostra com retiradas.

O gráfico 5.2, painel superior, mostra a distribuição da renda segundo a adesão ou não dos países na amostra com retiradas. A renda mediana dos países que aderiram é maior que a dos que não aderiram. Similarmente, o terceiro quartil dos países participantes do tratado é maior do que entre não participantes. Como o intervalo o primeiro quartil representa o ponto que contém 25% dos dados e o terceiro quartil 75% dos dados, o interquartil contém 50% dos dados. Assim, um terceiro quartil maior indica países com renda per capita mais alta entre os que aderiram.

Para checar a possibilidade de viés induzido pela decisão de retirar da amostra os estados após a adesão à Convenção de Paris, podemos comparar os gráficos superior e inferior (amostra sem retiradas) do gráfico 5.2. Como se pode observar, a tendência de quanto maior a renda per capita maior a proporção de adesão à Convenção de Paris induz algum viés negativo. De fato a renda per capita mediana na amostra com retiradas é de US\$ 7.470, enquanto que na amostra sem retiradas é de US\$ 9.010. Além disso, o terceiro quartil no gráfico do painel superior é de US\$ 13.733 e no gráfico do painel inferior, 19.562. Esses

dados sugerem que há uma tendência dos países de renda per capita mais alta aderirem cedo à Convenção, reduzindo assim a presença de tais estados na amostra. Portanto, temos um viés negativo, de subestimação, o que nos permite identificar limites inferiores.

Na amostra completa (figura 5.1a), a mediana amostral dos estados que aderem à Convenção de Paris é sempre maior do que a mediana amostral dos que não aderem. No painel direito (5.1b) vemos um padrão bastante similar em termos de distribuição das medianas, com exceção do ano 1985. Com relação aos quartis, que medem a variabilidade dos dados, vemos que, na amostra com retiradas de países, os quartis dos países que aderem são menos distantes entre si do que na amostra completa. Isto reflete o fato, apontado por nós, de que se a associação é positiva, os países com maiores valores da variável independente são retirados da amostra. É possível, portanto, que haja algum viés na estimação da associação entre renda e probabilidade de adesão. Porém, como esperávamos teoricamente que a associação fosse positiva e as estatísticas descritivas apontam nessa direção, é possível afirmar que o viés deve subestimar o efeito da associação positiva. É possível, portanto, identificar limites inferiores para o efeito da renda sobre a probabilidade de adesão na análise regressão.

### 5.3. Regime Político

Regime político é uma variável que procura medir se um país num dado ano é democracia ou autocracia. Nós utilizamos a variável compilada por Cheibub, Ghandi e Vreeland (2009), denominada **democracy**, que é uma atualização e aperfeiçoamento da codificação dicotômica introduzida por Przeworski, Alavarez, Cheibub e Limongi (2000). Os detalhes da codificação da variável podem ser encontrados em Cheibub, Ghandi e Vreeland (2009). Para os quinquênios, tomamos como observação o valor da variável no ano considerado.

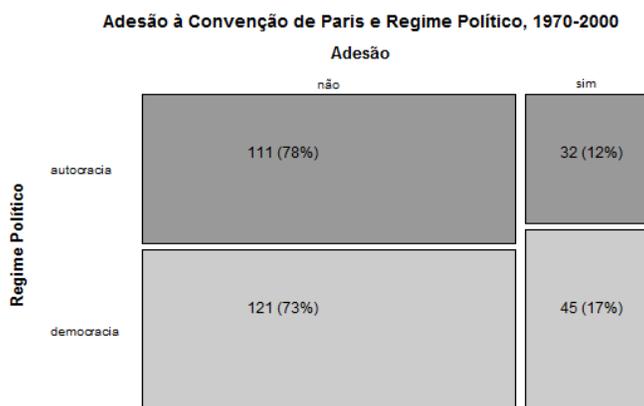


Gráfico 5.3. Mosaic Plot de regime político e adesão à Convenção de Paris. O Mosaic plot é uma representação gráfica das proporções de adesão por Regime Político. Assim, do total de autocracias da amostra, 78% não aderiram à Convenção, enquanto 22% aderiram. E do total de democracias, 73% aderiram, e 27% aderiram. Essa maior proporção de democracias que aderem à Convenção em relação a autocracias é visualizada no gráfico pelo fato da área do retângulo representando democracias que aderem (“sim”) é maior do que a área de autocracias que aderem. Por outro lado, a área de autocracias que não aderem (“não”) é maior do que a área de democracias que não aderem. Essa diferença, porém, não é estatisticamente significativa. Em outras palavras, nossa amostra não conseguiu detectar efeito de democracia sobre adesão.

O gráfico 5.3 apresenta um **mosaic plot**, que é uma representação gráfica da tabela de contingência. O gráfico mostra se regime político está associado a maior ou menor adesão. Os números apresentam o total de observações e entre parênteses o percentual de observações de cada regime. Assim, vemos que 78% das autocracias não aderem, e 22% aderem ao tratado. Entre países democráticos, 73% não aderem ao tratado, enquanto apenas 27% aderem. Essa diferença, porém, não é significativa ao nível de 10%, indicando que com nossa amostra não conseguimos detectar o efeito do regime político sobre a decisão de aderir ao tratado.

## 5.4. Nível Educacional

Para mensurar o nível educacional, utilizamos os dados disponíveis em Barro e Lee (2000). Em que pese alguns problemas de mensuração contidos nesse banco de dados (Portela, Alessie, Teulings, 2006), é um banco de dados clássico em economia e utilizado na literatura em artigos importantes (Glaeser, Ponzetto, Shleifer, 2007; Acemoglu, Johnson, Robinson e Yared, 2005). O banco contém várias medidas de escolaridade a cada 5 anos por países. Nós utilizamos o percentual da população com idade acima de 15 anos com escolaridade maior que nível secundário. Escolhemos essa variável porque entendemos que o conhecimento de

nível secundário ou superior é uma medida mais aproximada da capacidade da mão-de-obra contribuir para o progresso tecnológico por meio da absorção ou mesmo inovação (Teixeira e Fortuna, 2006; Barro e Lee, 2000).

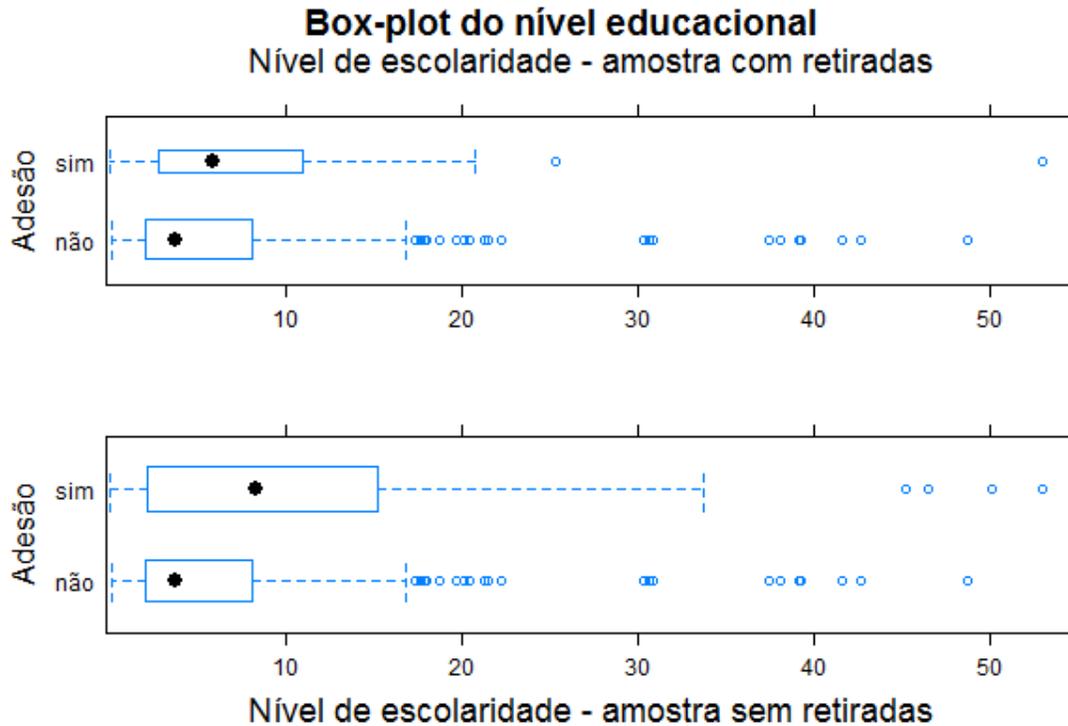


Gráfico 5.4. Box-plot da escolaridade (percentual da população com nível acima de secundário) entre países que aderiram (sim) e não aderiram (não). O gráfico mostra associação positiva entre educação e adesão. A educação mediana é maior entre os países que aderiram do que entre os que não aderiam. O mesmo é verdadeiro para o primeiro e terceiro quartil.

O gráfico 5.4 (painel superior) apresenta a distribuição do nível de educação entre países que aderiram e não aderiram à Convenção de Paris. Embora a mediana e o terceiro quartil dos países que aderem seja maior do que a dos estados não aderentes, temos vários *outliers* no extremo superior da distribuição entre os países que não aderem à convenção. Isso significa que a distribuição do nível de escolaridade entre não é aderentes é mais assimétrica do que entre estados aderentes, com cauda longa entre estados com altos níveis de escolaridade.

A comparação das amostras com retiradas (painel superior) e sem retiradas (painel inferior) apresenta bastante diferenças. O terceiro quartil que contém 75% dos dados, é muito maior na amostra sem retiradas do que na amostra com retiradas. Na amostra com retiradas, 75% das observações entre estados aderentes possuem nível de escolaridade menor do que 10,9% da população com nível maior que secundário. Já na amostra com retiradas, esse valor é de 15%.

Uma explicação para esse fato pode estar na decisão de excluir da amostra os países após eles aderirem ao tratado. Os EUA, por exemplo, que ratificaram o tratado em 1973, aparecem na amostra apenas nos anos 1970 e 1975. Assim, os países que continuam na amostra seriam aqueles com menor nível de escolaridade. Com o passar do tempo, a média amostral dos países que aderem ao tratado seria menor do que a média que observaríamos se todos os países permanecessem na amostra.

Isso sugere que há um viés induzido pelo nosso procedimento de retirar da amostra países após terem aderido ao tratado. Esse viés deve diminuir o efeito da educação sobre a probabilidade de adesão. Assim, nós temos novamente uma identificação dos limites inferiores, já que estaremos subestimando o efeito. Porém, não é possível estabelecer quantitativamente os limites superiores, pois não são identificáveis.

## **5.5. Legislação doméstica em patentes**

Conforme argumentado no segundo capítulo, utilizamos o índice Ginarte-Park, consagrado pela literatura, para mensuração do grau de proteção patentária dos países. Na análise estatística decidimos utilizar uma variável defasada em um quinquênio, pois esperamos que mudanças em períodos anteriores tivessem seus impactos apenas nos anos seguintes, e não no próprio ano. Os detalhes das modificações empreendidas por nós para tornar os dados adequados aos nossos propósitos encontram-se no referido capítulo. Vale destacar que nós *imputamos* os valores do Reino Unido para o período pré-1995, pois não tínhamos os dados desagregados para o período 1970-1990. Porém, como Park (2008) informa o valor médio do índice para esse período, decidimos repetir esse valor médio para os anos 1965-1990. Embora haja imprecisão nesse método, a maior parte dos países desenvolvidos não alterou sua legislação doméstica de patentes no período considerado, de forma que o índice para esses países é quase constante nesse período. Assim, não é uma aproximação tão ruim quanto pode parecer à primeira vista.

Os gráficos 5.5 apresenta a distribuição dos dados do índice Ginarte-Park para todo o período na amostra com retiradas e na amostra completa (sem retiradas).

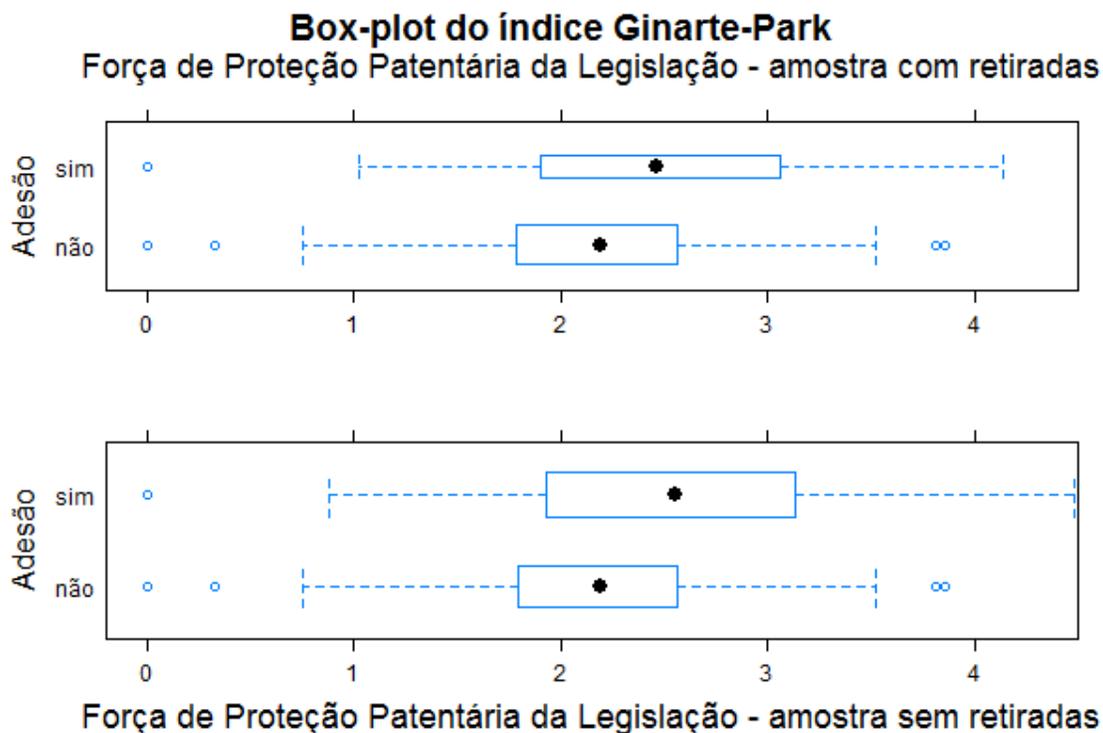


Gráfico 5.5. Box-plot do índice de força da proteção patentária da legislação doméstica, medido pelo índice Ginarte-Park na amostra com retiradas (retiramos os países da amostra após adesão), no painel superior, e sem retiradas (como seria a amostra se não adotássemos o procedimento de retirar os países da amostra após adesão), no painel inferior. A comparação dos dois painéis não apresenta grandes diferenças. Em ambos a associação é positiva e na amostra com retiradas a mediana é levemente menor do que na amostra com retiradas. Se o procedimento introduzir algum viés nessa variável, deve ser positivo.

Como se pode observar no gráfico 5.5, painel superior, que apresenta a distribuição dos dados da amostra utilizada nas regressões, os dados são bem distribuídos e com poucos *outliers*. Há uma diferença entre o índice Ginarte-Park dos países que aderem e não aderem. A força da proteção de patentes na legislação doméstica é maior entre estados aderentes da Convenção de Paris do que entre não aderentes. Assim, não condicional a outros fatores, estados com maior proteção de patentes aderem mais à Convenção.

Ao compararmos a amostra com retiradas (painel superior) e sem retiradas (painel inferior), vemos que há pouca diferença, indicando que o procedimento não introduz vieses substantivos. Entretanto, seguindo a abordagem de Manski (2009), se houver algum viés decorrente do procedimento, este é positivo e, portanto, torna os limites inferiores dos intervalos de credibilidade identificados.

## 5.6. Patentes nos Estados Unidos

Como *proxy* para a importância de obter proteção patentária no estrangeiro, utilizamos o número de patentes obtidas nos EUA. Os dados foram retirados do website do *US Patent Trade Office* (USPTO, 2009). Para cálculo dos dados quinquenais, utilizamos a média de cinco anos, similarmente à renda per capita.

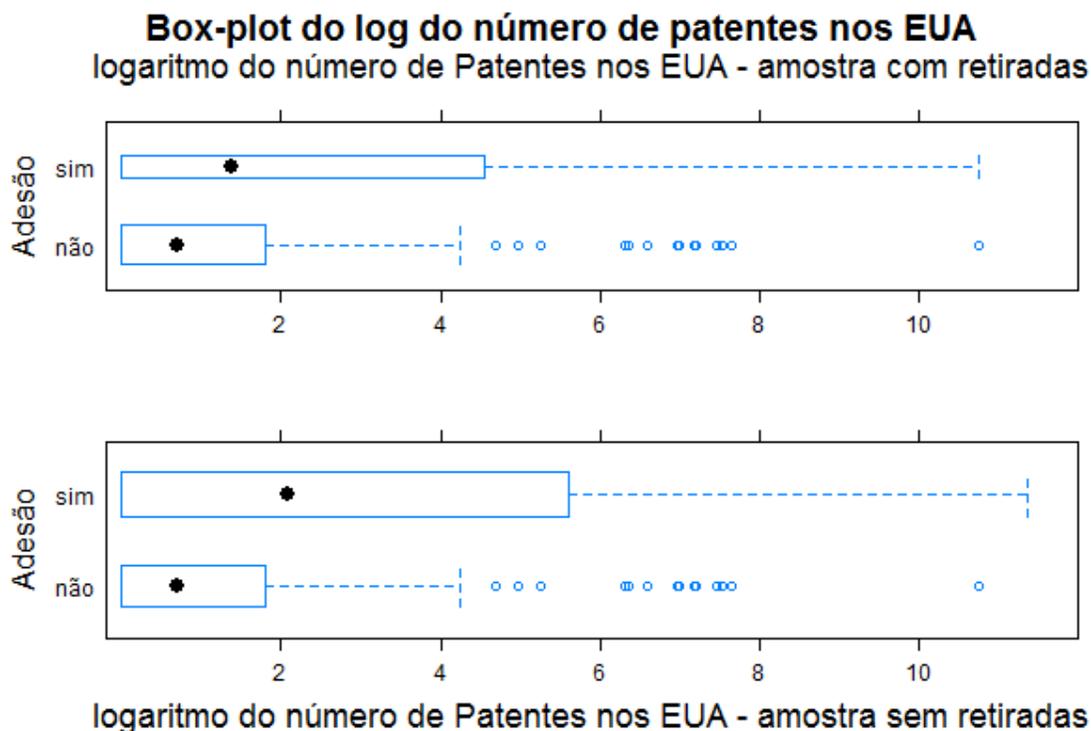


Gráfico 5.6. Box-plot do logaritmo do número de patentes nos EUA na amostra com retiradas (retiramos os países da amostra após adesão), no painel superior, e sem retiradas (como seria a amostra se não adotássemos o procedimento de retirar os países da amostra após adesão), no painel inferior. A comparação dos dois painéis não apresenta grandes diferenças. Em ambos a associação é positiva e na amostra com retiradas a mediana é um pouco menor do que na amostra com retiradas. Se o procedimento introduzir algum viés nessa variável, deve ser positivo.

O gráfico 5.6 (painel superior) apresenta a distribuição dos dados agregados do logaritmo do número de patentes nos EUA entre países que aderiram e não aderiram à Convenção de Paris. Para cômputo do logaritmo natural do número de patentes, adicionamos 1 ao número de patentes de todos os países, para evitar problemas com países com zero patentes. Assim, os dados representam o  $\log(\text{número de Patentes país}_{i,t} + 1)$ .

A mediana dos países que aderiram é maior do que a dos que não aderiram. Como há um grande número de países com zero patentes, a diferença nos quartis se dá no terceiro quartil,

onde se vê que as patentes obtidas no exterior pelos países que aderem são mais dispersas no extremo superior da distribuição.

Analisando os mesmos na amostra sem retiradas de países (painel inferior), vemos que a mediana é um pouco maior entre os países que aderem no painel inferior (amostra sem retiradas). Ademais, o terceiro quartil entre os países que aderiram ao tratado da amostra com retiradas é menor do que na amostra sem retiradas. Assim, se houver algum viés devido ao nosso procedimento, ele deve ser positivo e permite identificar limites inferiores para o parâmetro de interesse.

## **5.7 Investimento Externo Direto (IED)**

Os dados de investimento externo direto (IED) foram retirados do bando de dados do Banco Mundial (2010). Os dados de investimento externo direto refletem o fluxo líquido de investimentos adquiridos pelo menos 10% do capital votante de empresas operando em outro país que não o do investidor Banco Mundial (2010). Pode ser pensado também como a soma do *equity capital*, reinvestimento de lucros e outros capitais de longo e curto prazo tal como mostrado nos Balanços de Pagamento dos Países. Os dados são em dólares correntes. Assim, dados de IED podem tanto ser positivos como negativos. IED negativo significa que houve mais desinvestimento externo do que investimento externo.

Como o tamanho da economia de um país afeta o montante de IED que ela pode receber, parece-nos mais conveniente utilizar a proporção de IED em relação ao PIB como medida do fluxo de investimento externo direto. Para obter os dados de cada quinquênio, nós dividimos a média dos cinco anos do fluxo de investimento externo direto pela média do PIB do país no mesmo período, medido em poder de paridade de Compra. Testes de robustez foram feitos para variações na forma de medir essa variável (por exemplo, dividindo pelo PIB a preços correntes) e são apresentados no anexo.

Para visualizar a dispersão dos dados de investimento externo direto, decidimos utilizar o histograma dos dados.

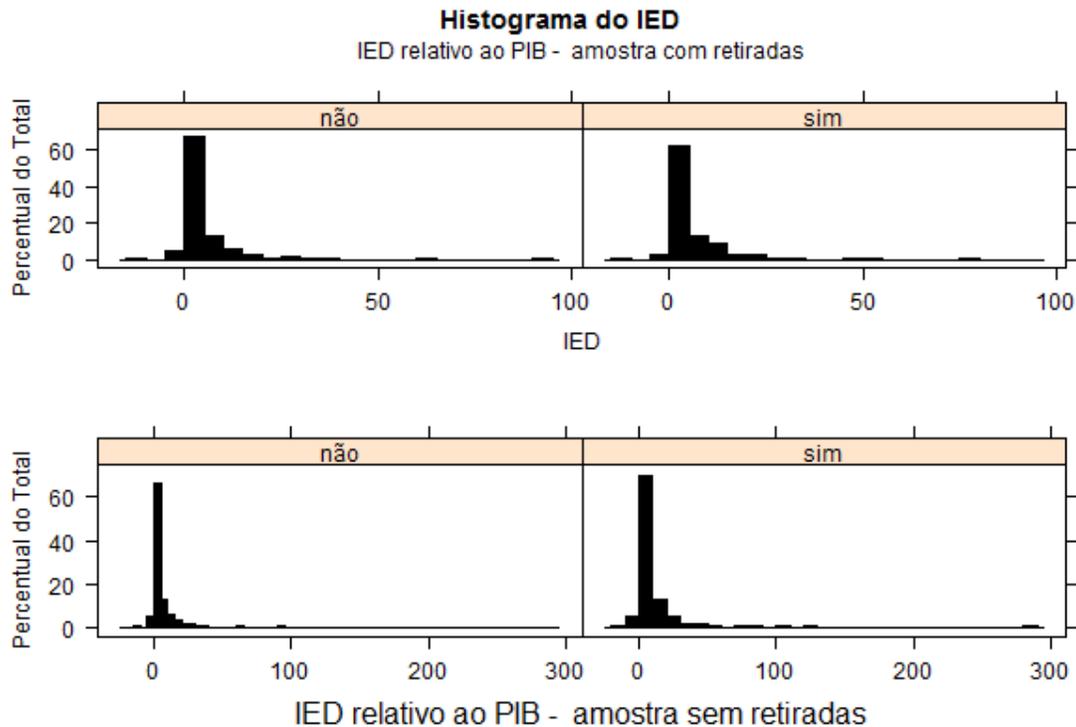


Gráfico 5.7. Histograma para o Investimento Externo Direto em relação ao PIB entre países que aderiram (sim) e não aderiram (não) à Convenção, na amostra com retiradas (painel superior) e sem retiradas (painel inferior). No painel superior, vemos que praticamente não há diferença na distribuição entre estados aderentes e não aderentes ao tratado. O valor modal é próximo do zero, indicando baixo fluxo de investimento externo direto. Na comparação entre amostras (painel superior e inferior), a principal diferença está em que alguns *outliers* que aderiram (painel inferior direita) não estão presentes no painel superior direita.

O gráfico 5.7 mostra o histograma para o percentual de investimento externo direto em relação ao PIB. O valor modal é próximo de 5% e há alguns valores extremados com investimento externo direto maior do que o PIB. A razão para tanto é que o IED foi mensurado em dólares correntes, enquanto que o PIB está mensurado em dólares constantes. Como não tínhamos deflator interno do PIB para todos os países em todos os anos, usamos essa medida de forma aproximada para ter o peso do IED em relação ao PIB dos países.

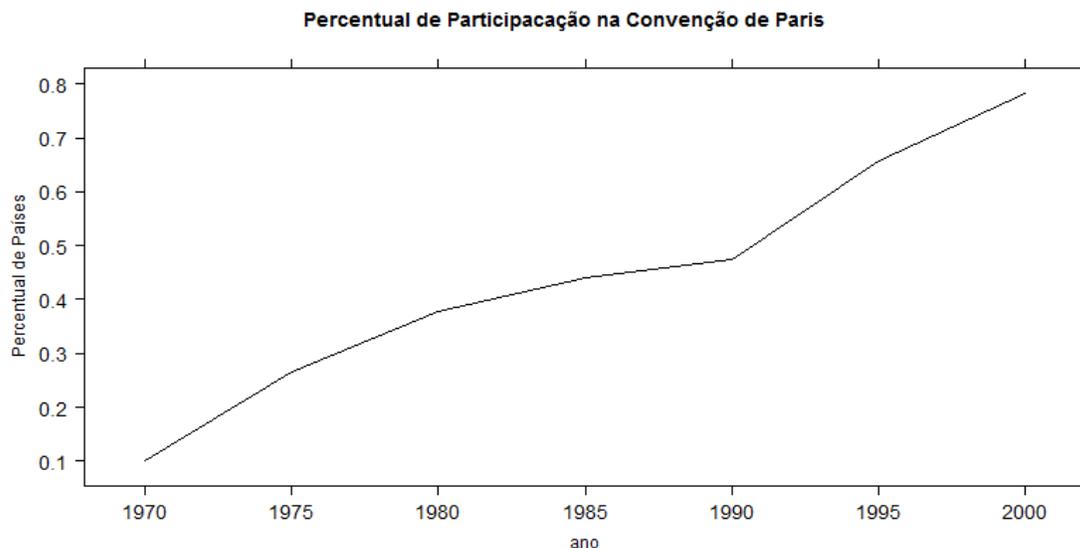
Os gráficos mostram um histograma para a amostra sem e com retiradas de países, respectivamente. Ao contrário dos outros gráficos, porém, não percebemos diferenças substanciais de uma amostra para outra, com exceção de alguns *outliers*. Assim, nesse caso não é possível identificar viés na estimação da variável.

## 5.8. Variáveis ao nível do quinquênio

Como estamos estimando um modelo hierárquico de dois níveis com intercepto variando por quinquênio, incluímos como preditor dos interceptos o percentual de países participantes da

Convenção de Paris por quinquênio. Para tanto, calculamos, pra cada ano, o percentual de países que aderiram ao tratado em relação ao total de países que poderiam ter aderido. Nessa variável foram incluídos mais do que os 90 países do primeiro nível, pois tínhamos informações para todos os países da base de dados se havia aderido ou não ao tratado. O número total de países, porém, varia no tempo, já que nesse período há tanto a criação como dissolução de estados independentes.

O gráfico 5.8 apresenta a evolução da participação na Convenção de Paris. Como se vê, há uma tendência crescente no tempo de adesão. Na estimação do modelo, nós utilizamos a variável defasada em cinco anos, de forma evitar problema de causalidade reversa, já que a adesão num dado ano aumenta a participação daquele ano, criando uma correlação espúria.



**Gráfico 5.8.** Percentual de estados membros da Convenção de Paris por ano. No começo dos anos 70, menos de 10% dos países considerados participavam do tratado. Ao final dos anos 2000, 80% dos estados da amostra haviam aderido ao acordo.

## 5.9. Regressão

Nós estimamos uma regressão logística com intercepto variando por ano. Um modelo com intercepto variando por país também foi testado, porém o modelo com intercepto variando por ano teve um ajuste superior. O anexo do trabalho apresenta alguns resultados do modelo com intercepto variando por país. O gráfico 5.1 abaixo apresenta os resultados do modelo estimado (os interceptos variando por país são mostrados na próxima seção). Os modelos estimados foram:

Modelo 1, “modelo.time”:

$$Y_{it} = \text{Bernoulli}(p_{it}),$$

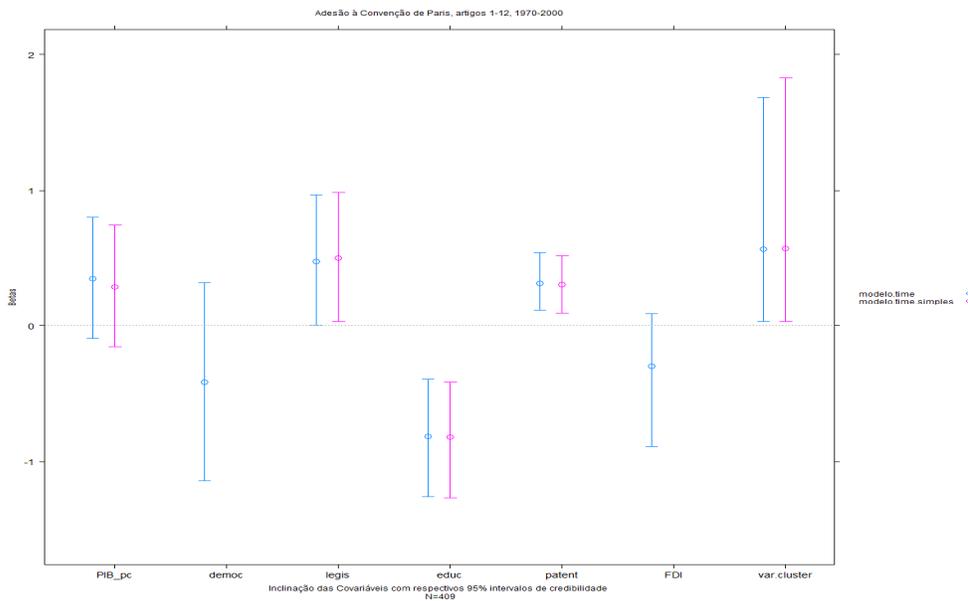
em que:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{\text{renda}} * \text{renda}_{it} + b_{\text{democracia}} * \text{democracia}_{it} + b_{\text{educ}} * \text{educ}_{it} + b_{\text{patente}} * \text{patente}_{it} + b_{\text{legis}} * \text{legis}_{it} + b_{\text{fdi}} * \text{fdi}_{it})$$

Para o segundo modelo, “modelo.time.simples”:

$$Y_{it} = \text{Bernoulli}(p_{it}),$$

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{\text{renda}} * \text{renda}_{it} + b_{\text{educ}} * \text{educ}_{it} + b_{\text{patente}} * \text{patente}_{it} + b_{\text{legis}} * \text{legis}_{it})$$



**Gráfico 5.9.** O gráfico apresenta os resultados de duas regressões Bayesianas hierárquicas no primeiro nível (nível dos países). Os pontos representam o efeito médio e as barras o intervalo de credibilidade a 95%. A linha pontilhada cruza o zero e facilita a visualização da significância estatística. Barras que não cruzam a linha pontilhadas indicam variáveis significativas. O modelo 1, denominado “modelo.time”, em azul, inclui as variáveis: renda per capita (“PIB\_pc”), democracia (“democ”), legislação patentária (“legis”), nível educacional (“educ”), logaritmo do número de patentes nos EUA (“patent”), investimento externo direto como percentual do PIB (“fdi”) e a variância dos interceptos (“var.cluster”). O modelo 2, denominado “modelo.time.simples”, excluiu da regressão as variáveis não significativas.

O gráfico 5.9 mostra os resultados da regressão multinível de dois modelos. O primeiro modelo – em rosa – é o resultado da regressão com intercepto variando por quinquênio e com todas as variáveis testadas por nós. O segundo modelo também é uma regressão com intercepto variando por período, porém, sem a inclusão das covariáveis “democracia”,

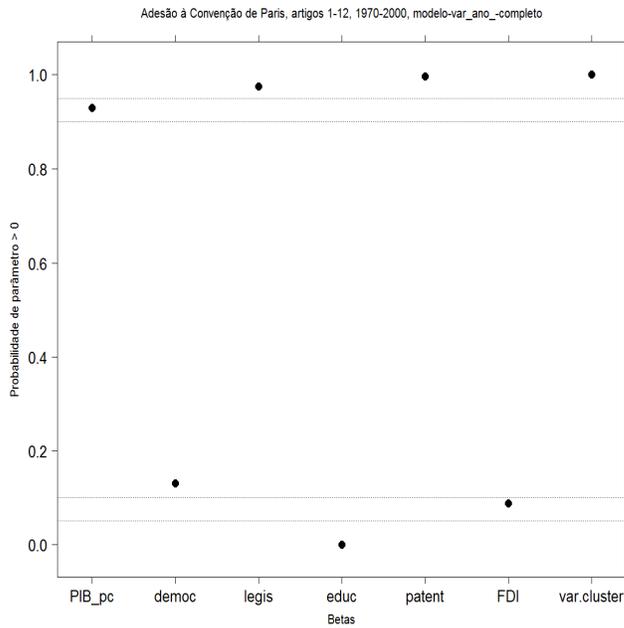
representada no gráfico por “democ” e IED, representada no gráfico por “FDI”. As barras representam 95% de intervalos de credibilidade.

Como se pode observar, educação, legislação doméstica de patentes e patentes nos EUA são significativas em ambos os modelos. Contudo, o nível educacional tem o “sinal trocado”, isto é, quanto maior o nível de escolaridade de um país, menor a probabilidade de adesão à Convenção de Paris, contrariamente ao esperado inicialmente.

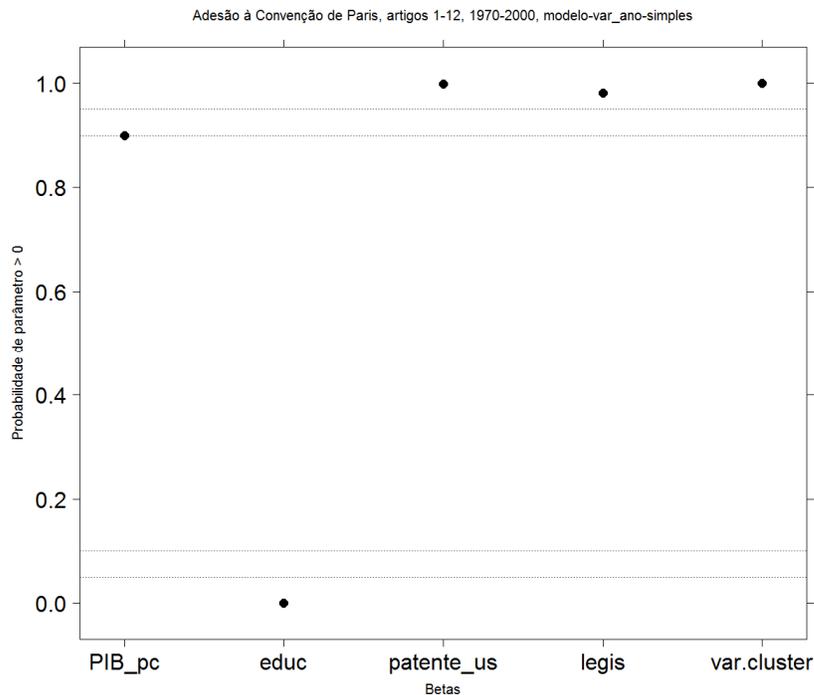
Outra forma de apresentar os resultados é nos perguntarmos a probabilidade de que um parâmetro seja maior do que zero (ou menor do que zero). No paradigma tradicional de Neyman-Pearson (Gelman e Tuerlinckx, 2000), a tomada de decisão é realizada tendo por base dois tipos de erros que se podem cometer: erro do tipo I, isto é, afirmar que a hipótese nula é falsa quando ela é verdadeira, e erro do tipo II, afirmar que a hipótese alternativa é falsa quando é de fato verdadeira. O erro do tipo I pode ser pensado como um falso positivo e erro do tipo II falso negativo (falha em detectar o efeito).

De acordo com Gelman e Tuerlinckx (2000), os erros mais importantes são o erro do tipo S (erros de Sinais) e erros do tipo M (erros de Magnitude). Um erro do tipo S ocorre quando nós concluimos que um efeito é positivo (negativo) quando na verdade ele é negativo (positivo). Em outras palavras, quando nós somos levados a concluir que um parâmetro é maior (menor) que zero, quando na verdade é menor (maior) que zero. E erros de tipo M, ou de magnitude, ocorrem quando nós afirmamos que o tamanho do efeito é maior (menor) do que ele realmente é. Assim, vejamos as probabilidades de que o sinal das variáveis seja positivo.

Os gráficos 5.9a e 5.9b mostram as probabilidades de que os parâmetros estimados sejam positivos.



**Gráfico 5.9a.** O gráfico apresenta as probabilidades a posteriori para cada parâmetro de serem maior do que zero, tal como estimados no modelo 1. As linhas pontilhadas indicam 95%, 90%, 10% e 5%. Renda per capita é maior que zero com mais de 90% de certeza, legislação, patentes nos EUA e “var.cluster” são maiores que zero com mais de 95% de certeza. Educação é maior que zero com menos de 5% de certeza, ou seja, é menor do que zero (negativo) com mais de 95% de certeza. Regime não é significativo e IED (“FDI”) é negativo com 90% de certeza.



**Gráfico 5.9b.** O gráfico apresenta as probabilidades a posteriori para cada parâmetro de serem maior do que zero, tal como estimados no modelo 2. As linhas pontilhadas indicam 95%, 90%, 10% e 5%. Renda per capita é maior que zero com mais de 90% de certeza, legislação, patentes nos EUA e “var.cluster” são maiores que zero com mais de 95% de certeza. Educação é maior que zero com menos de 5% de certeza, ou seja, é menor do que zero (negativo) com mais de 95% de certeza. Regime não é significativo e IED (“FDI”) é negativo com 90% de certeza.

As linhas pontilhadas refletem, de cima para baixo, a probabilidade de o parâmetro ser maior do que zero a 95%, 90%, 10% e 5% respectivamente. Assim, legislação, patentes e variância do intercepto têm mais de 95% de chance de serem maior do que zero. Renda per capita tem mais de 90% de probabilidade de ser maior que zero. Educação tem menos de 5% de chance de ser maior do que zero, isto é, mais de 95% de chance de ser negativo, FDI tem menos de 10% de probabilidade de ser positivo (logo, mais de 90% de chance de ser negativo) e democracia tem mais de 10% de chance de ser positivo. Como algumas variáveis não foram significativas, nós estimamos o modelo com intercepto variando por país sem algumas das covariáveis (legislação doméstica e IED). Renda per capita e patentes nos EUA continuaram com probabilidades a posteriori de serem positiva superiores a 95%, e educação com probabilidade inferior a 5%.

Como vimos, quer olhemos para os intervalos de credibilidade, quer para a probabilidade de o parâmetro ser positivo, patentes nos EUA, legislação doméstica e nível educacional são consistentemente significativos. Renda per capita e IED são marginalmente significativos, e regime político não é significativo em nenhum modelo.

Esses resultados mostram, primeiramente, que variáveis diretamente relacionadas ao tema do tratado, como logaritmo do número de patentes nos EUA e a força da proteção patentária doméstica estão associadas positivamente com adesão à Convenção de Paris. Esses resultados são esperados dentro da teoria institucionalista-racionalista, segundo a qual são os custos e benefícios de um tratado que determinam se um país irá aderir ou não ao tratado.

Países com maior número de patentes no exterior têm mais a ganhar com a participação no tratado, pois a Convenção de Paris regula o direito de prioridade, permite ao país fazer parte do Tratado de Cooperação de Patentes e garante a cláusula de tratamento nacional, que é importante para evitar discriminação contra estrangeiros. Assim, quem mais tem a se beneficiar com o tratado nesses aspectos tem maior probabilidade de adesão.

Similarmente, países com legislação patentária mais forte no quinquênio anterior possuem menor custo de adesão à convenção, tanto em termos de *compliance* quanto em termos de necessidade de modificação da legislação. Assim, quanto menor o custo da adesão para o país, maior a probabilidade dele ter aderido ao tratado.

Nesse aspecto, vale notar que os achados sobre a importância do regime político da literatura sobre *compliance* e adesão em direitos humanos não são corroborados no caso da Convenção

de Paris. Não há diferenças estatisticamente significativas entre democracias e autocracias. Além disso, nós estimamos também um modelo com interação entre força da legislação doméstica e regime político do país, para checar a hipótese de que democracias com maiores custos de *compliance* adeririam menos que a contraparte autocrática. Não houve significância estatística nos coeficientes e esse resultado é reportado no anexo.

Renda per capita e investimento externo direto são marginalmente significativos, ou seja, apenas ao nível de 90%. Esses resultados apontam para evidências fracas da importância dessas variáveis. No caso da renda, nós estimamos também logaritmo da renda, porém este não foi significativo nem a 90%. Esses resultados sugerem, portanto, que com nosso tamanho amostral não conseguimos detectar o efeito dessas variáveis. Em outras palavras, como o efeito não foi detectado de forma robusta, é possível que esse efeito seja pequeno e por isso nossa amostra não tenha sido capaz de detectá-lo. De todo modo, a renda teve o sinal esperado (positivo) em todos os modelos. Esse achado está em acordo com a teoria segundo a qual países mais desenvolvidos teriam mais a ganhar com a Convenção de Paris e, portanto, maior taxa de adesão.

O IED teve sinal negativo, o que sugere um efeito pequeno de que, condicional às demais covariáveis, países com menor fluxo de investimento externo direto em média adeririam mais à Convenção de Paris. Uma possível explicação para esse achado é que a relação de causalidade pode ser inversa. Países que aderem à Convenção de Paris recebem mais investimento externo direto e, portanto, países com menor fluxo de investimento teriam a expectativa de obter um benefício colateral maior da adesão à Convenção. Porém, devido a esse problema de causalidade inversa e efeito pouco robusto, não é possível afirmar que temos evidências fortes num sentido ou noutro.

## **EDUCAÇÃO**

A variável nível educacional, utilizada como *proxy* do estoque de capital humano do país apresenta o resultado mais contra-intuitivo. Países com menor nível de escolaridade da população participam mais da Convenção de Paris do que países com maior nível educacional.

Uma possibilidade explicativa está em viés de variável omitida ou erros de mensuração. A literatura sobre capital humano em perspectiva comparada discute há algum tempo os problemas de mensuração de estoque de educação dos países, especialmente nos países

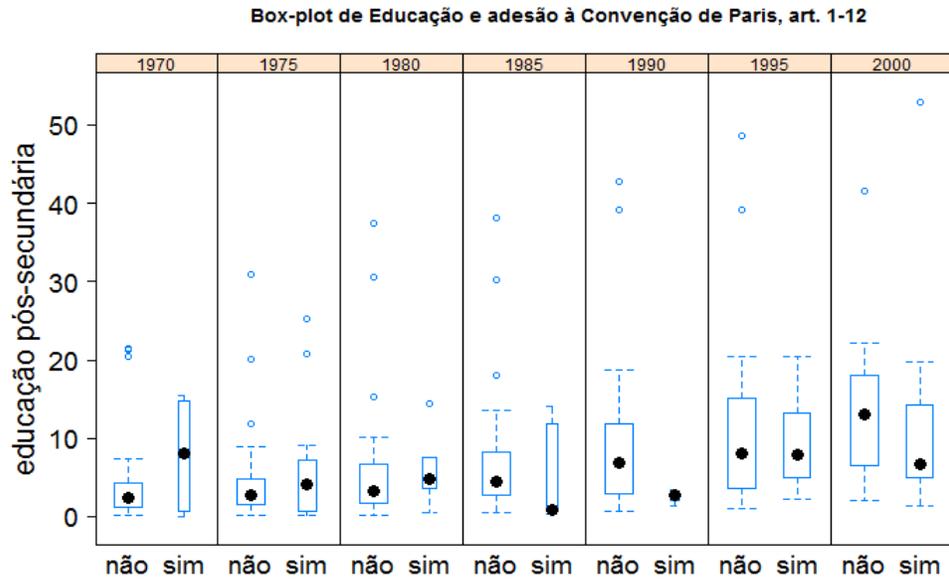
menos desenvolvidos. Dados são ruins e é necessário interpolação para alguns anos e países, mesmo em se tratando de dados quinquenais (Barro e Lee, 2000).

Outra possibilidade aventada por nós é que escolaridade e renda per capita estivessem medindo aspectos similares da economia de um país, pois países mais desenvolvidos não somente têm maior renda per capita como também possuem maior nível educacional. De fato, a correlação entre renda per capita e educação é de 60% na nossa amostra. Porém, ao estimarmos um modelo sem renda per capita o sinal da variável educação é o mesmo, de forma que excluímos essa possibilidade.

Outra explicação possível é que, na nossa amostra, países recém-independentes aderem cedo à Convenção de Paris em relação à data de independência. E esses países poderiam ter um nível educacional mais baixo, o que estaria viesando o resultado. Exemplos desse comportamento são Malawi e Senegal, que aderiram ainda nos anos 70 à Convenção. Porém, a inclusão de uma variável idade do regime não foi significativa e não alterou o sinal da variável educação.

Outra possibilidade é, seguindo a teoria institucionalista, segundo a qual são as expectativas de benefícios que influenciam as decisões dos países. Assim, países com menor nível educacional poderiam esperar um benefício colateral maior (em termos de atração de investimentos externos diretos, aumentos nos retornos à educação domesticamente) de participar da Convenção de Paris do que países com maior nível de escolaridade. Essa explicação é um pouco mais plausível se levarmos em conta que IED também teve sinal negativo, embora não significativo a 95%.

Para entender quais aspectos dos nossos dados estão conduzindo a esses resultados, decidimos voltar então para a estatística descritiva dessa variável.



**Gráfico 5.9.1.** Box-plot do nível de escolaridade por ano, entre países que aderiram (“sim”) e não aderiram (“não”) à Convenção de Paris. É possível observar que para o ano de 1970, a mediana dos estados aderentes é muito maior do que a dos estados não aderentes. Essa relação é variável por ano.

O gráfico 5.9.1 apresenta, para cada quinquênio, um box-plot do nível de escolaridade entre os países que aderiram e não aderiram à Convenção de Paris. Na maior parte dos quinquênios não há diferença entre adesão ou não adesão, porém é possível que haja alguma diferença antes e após os anos 80. Até os anos 80 países com maior nível de escolaridade aderem mais ao tratado do que países com menor nível educacional, embora o número de observações de países que aderiram ao tratado nesse período seja pequeno. Após os anos 80, especialmente em 1985 e 1990 e 2000 há uma inversão na associação entre educação e adesão.

Surgiu então a hipótese de que o efeito da educação podia variar por ano, de forma que seria interessante estimar um modelo multinível com inclinação da educação variando por ano. Como os dados de adesão são poucos, a estimação Bayesiana multinível, ao fazer o *partial pooling* ou *shrinkage* dos dados em direção à média comum, permite-nos estimar um modelo desse tipo com menor incerteza. Contudo, os resultados são essencialmente os mesmos, ou seja, educação tem um efeito negativo em todos os anos e a variância das inclinações é muito

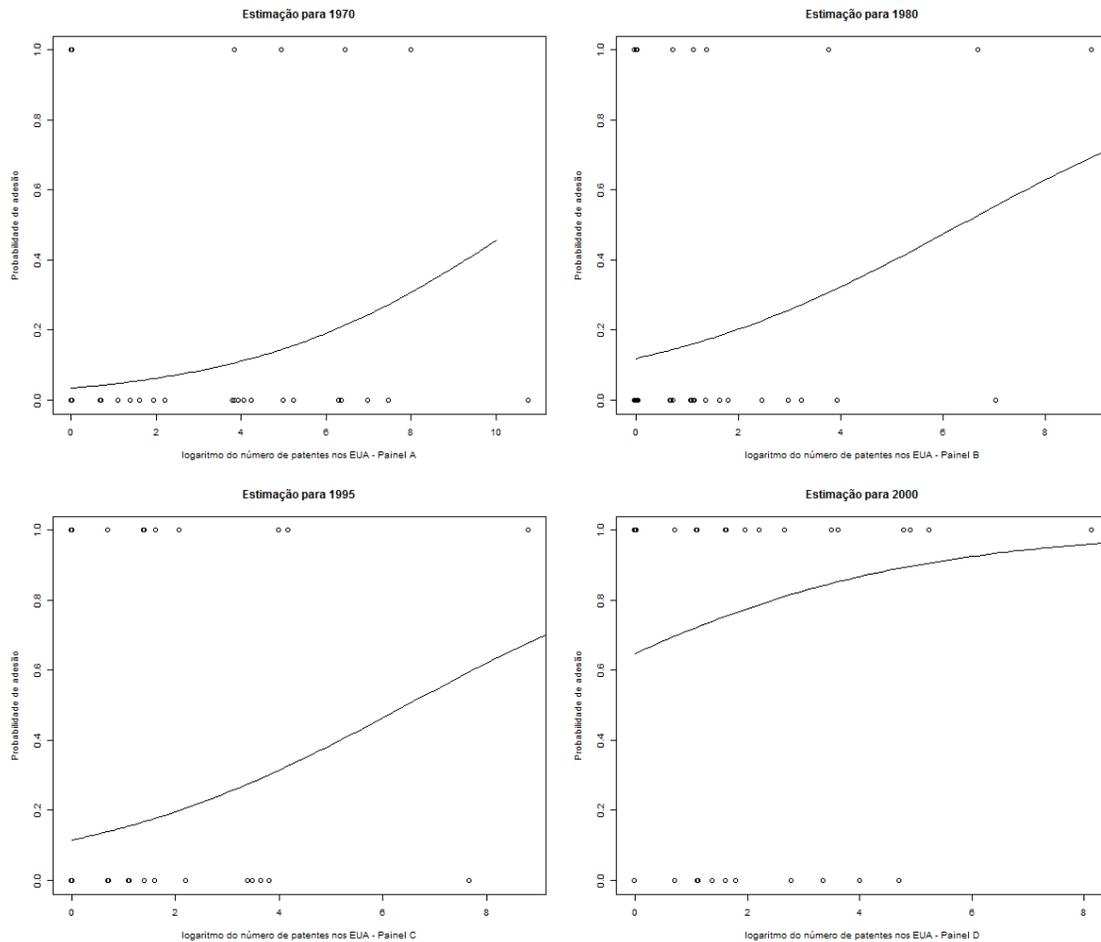
pequena e não é significativamente diferente do zero, o que implica que não temos motivo para supor que as inclinações sejam diferentes. Além disso, o ajuste do modelo, medido pelo DIC – *Deviance Information Criteria*, não foi melhor do que o modelo mais simples sem variação na inclinação.

DIC - *Deviance Information Criterion* – é uma medida de ajuste do modelo estimado, definido por  $D_{hat} + 2P_d$ , em que  $D_{hat}$  é uma estimativa pontual do desvio (*deviance*) utilizando os valores médios da posteriori dos parâmetros. E  $P_d$  é uma medida do número efetivo de parâmetros. Assim, o DIC é similar ao AIC, *Akaike Information Criterion*, já que permite comparar modelos com a mesma base amostral, mas não a veracidade do modelo em si. Além disso, é mais adequado que o AIC para modelos multiníveis, já que AIC não leva em consideração os parâmetros de nível superiores (termos aleatórios). De acordo com Spiegelhalter, Best, Carlin e Van der Linde (2002). Diferenças de mais de 10 no DIC praticamente deixam o modelo com DIC maior como sendo inferior, diferenças entre 5 e 10 são substanciais e diferenças menores que 5 não são muito confiáveis. Porém, como toda medida de resumo de ajuste do modelo, não é muito seguro confiar-se excessivamente em escolher modelos baseados no DIC.

Por fim, é possível que o efeito não seja negativo, já que se houver viés na análise introduzido pela exclusão dos estados da amostra após adesão ao tratado, este viés é negativo. Assim, não é possível identificar limites superiores, de forma que um resultado significativo pode não ser negativo. Consideramos portanto as evidências não tão robustas.

## **MAGNITUDE DOS EFEITOS**

Pra complementar a interpretação dos resultados, avaliemos a magnitude dos efeitos. Essa análise é particularmente importante porque regressões logísticas são mais difíceis de interpretar, mesmo em modelos não-hierárquicos (Brambor, Clark, Golder, 2006; Flom, Strauss, 2003).



**Gráficos 5.9.2a, 5.9.2b, 5.9.2c, 5.9.2d. Os gráficos apresentam o efeito médio do logaritmo do número de patentes nos EUA em anos selecionados, para o modelo 2 (com exclusão das variáveis não-significativas).**

Os quatro painéis (gráfico 5.9.2a, 5.9.2b, 5.9.2c, 5.9.2d) mostram o efeito do logaritmo do número de patentes sobre a probabilidade de adesão em autocracias, tomando valores médios das demais variáveis. A curva nos gráficos mostra como a probabilidade de adesão (no eixo y) varia para cada nível do logaritmo do número de patentes (eixo x). O painel A mostra o efeito estimado para o ano de 1970, o painel B para o ano de 1980, o painel C para o ano de 1995 e o painel D para o ano 2000. Os círculos representam adesão ou não dos países. Círculos no alto de cada gráfico representam adesão, círculos na parte de baixo de cada gráfico representam não adesão.

Pela comparação dos gráficos e dos dados brutos, percebe-se que o efeito do número de patentes varia entre os países a depender do período estudado. No painel A, onde temos as estimativas para os anos 70, o efeito preditivo esperado de um aumento no número de patentes nos EUA é maior para os países do extremo superior da distribuição. Dizendo de

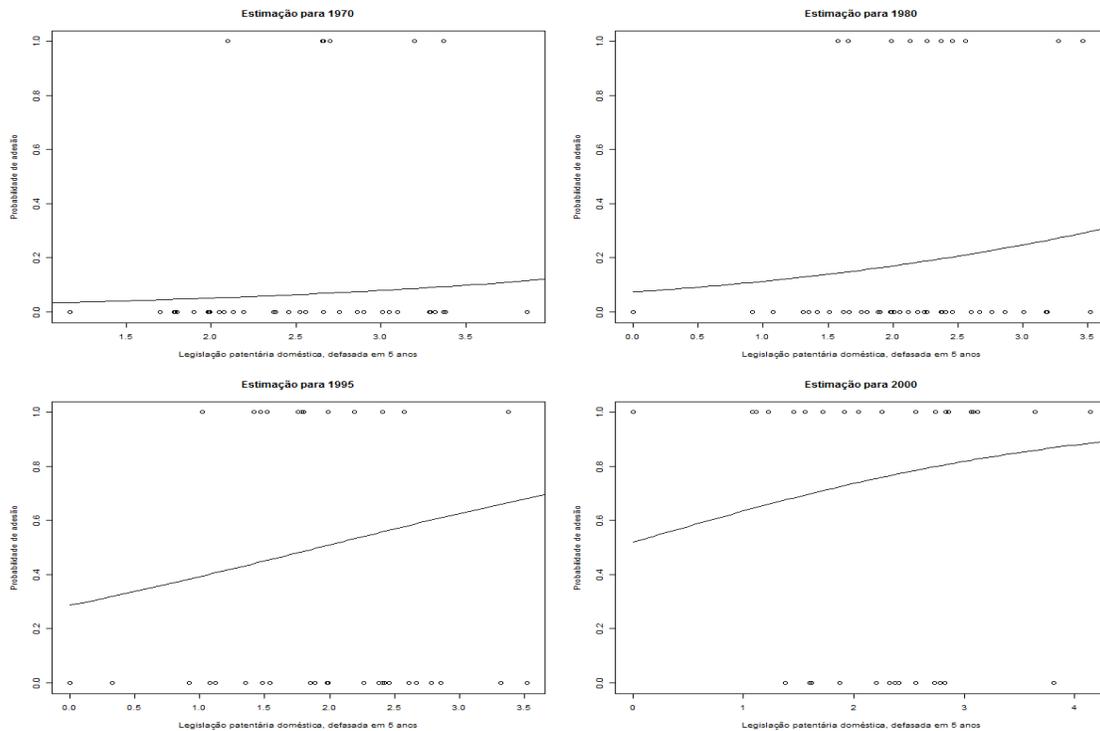
outro modo, países com muitas patentes nos EUA anos 70 que aumentassem seu número de patentes neste país teriam um probabilidade muito maior de adesão. Por exemplo, aumentos de 6 para 8 no logaritmo do número de patentes (de 400 para quase 3000 patentes) representam uma diferença preditiva de aproximadamente 10%, isto é, esse país aumentaria a probabilidade de adesão em 10%.

Contrariamente, variações no número de patentes nos EUA em países com poucas patentes têm pouco efeito preditivo nesse período. Apenas ilustrativamente, um país que passasse de zero patentes para 8 patentes nos EUA, um aumento expressivo, teria sua probabilidade de adesão aumentada em menos de 3%.

Por outro lado, o quadro é oposto nos anos 2000. A diferença preditiva entre um país sem patentes nos EUA e um país com 8 patentes nos EUA é de 15% nos anos 2000. Enquanto a diferença preditiva entre um país com 400 e outro com 3000 patentes nos EUA é de 3%.

Esse resultado é coerente com o que sabemos da evolução da arena internacional nos últimos 30 anos e também com a importância das patentes no cenário econômico internacional. Nas últimas décadas, mais e mais países procuram incentivar a obtenção de patentes no exterior, em particular nos EUA. Assim, há uma tendência de se aumentar o número de patentes obtidas nos EUA nas últimas décadas, de forma que obter 10 patentes em 1970 é diferente de obter 10 patentes nos anos 2000.

A mesma análise pode ser feita para as outras variáveis. Apresentamos aqui os resultados para legislação patentária doméstica:



Gráficos 5.9.3a, 5.9.3b, 5.9.3c, 5.9.3d. Os gráficos apresentam o efeito médio da força de proteção de patentes pela legislação doméstica, medida pelo índice Ginarte-Park, em anos selecionados, para o modelo 2 (com exclusão das variáveis não-significativas).

Como se pode observar pelos gráficos 5.9.3a, 5.9.3b, 5.9.3c, 5.9.3d, o efeito preditivo da variação da legislação patentária sob a probabilidade de adesão é praticamente nulo no começo dos anos 70. Comparativamente ao número de patentes nos EUA, portanto, o efeito preditivo esperado da variação na legislação doméstica de patentes é menor, particularmente nos anos iniciais de adesão ao tratado. Dito de outro modo, o número de patentes no exterior é variável mais importante para prever adesão à Convenção do que legislação doméstica em patentes.

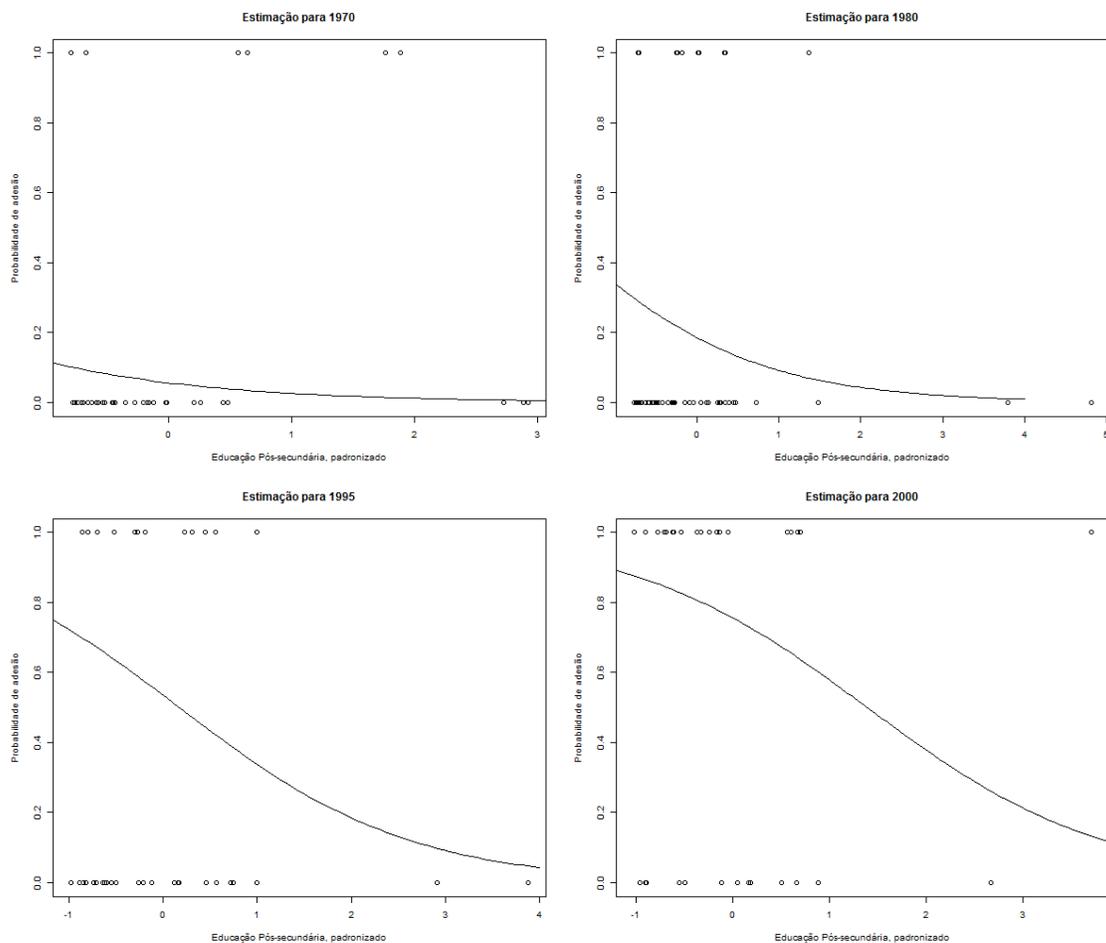
Com relação à variação do efeito por quinquênio, os resultados indicam que é após o fim da guerra-fria e criação da OMC que o efeito da legislação de patentes é mais importante na nossa amostra. De fato, pelos gráficos, é possível perceber que o efeito médio é maior após 1995. Uma diferença entre um país sem proteção de patente (índice 0) em 1995 e um país com índice 2 (tudo o mais constante) é de aproximadamente 20% na probabilidade de adesão.

Notem porém que essa diferença temporal não implica em dizer que a inclinação é diferente em 1995 e 1970, por exemplo. A diferença decorre do caráter não-linear da regressão logística e dos interceptos que variam por ano. O que o gráfico nos permite interpretar é que

no intervalo de dados relevantes em nossa amostra (com índices de proteção patentária variando entre 0 e 5), o efeito é maior em 1995 e 2000.

De todo modo, esses achados mostram que é quando a participação aumenta (ao final dos anos 90) que o grau de proteção patentária doméstica é mais importante como preditor da adesão. Em outras palavras, nos anos iniciais do tratado, essa variável não é um preditor relevante.

Por fim, apresentamos os mesmos painéis para a variável educação no gráfico 5.9.4a, 5.9.4b, 5.9.4c, 5.9.4d,.



Gráficos 5.9.4a, 5.9.4b, 5.9.4c, 5.9.4d. Os gráficos apresentam o efeito médio nível educacional em anos selecionados, para o modelo 2 (com exclusão das variáveis não-significativas).

Como se pode perceber, já a partir dos anos 80 o efeito preditivo de educação relevante. Como a maior parte dos dados está concentrada em torno da média (quase todos os pontos

estão no intervalo de mais ou menos um desvio padrão da média), o efeito preditivo estimado é relativamente grande. De fato a diferença preditiva esperada nos anos 80 entre um país com nível de educação resulta em um desvio-padrão abaixo da média e um país exatamente na média é de pouco mais de 10% em termos absolutos, isto é, espera-se que a probabilidade de adesão diminua em aproximadamente 10%.

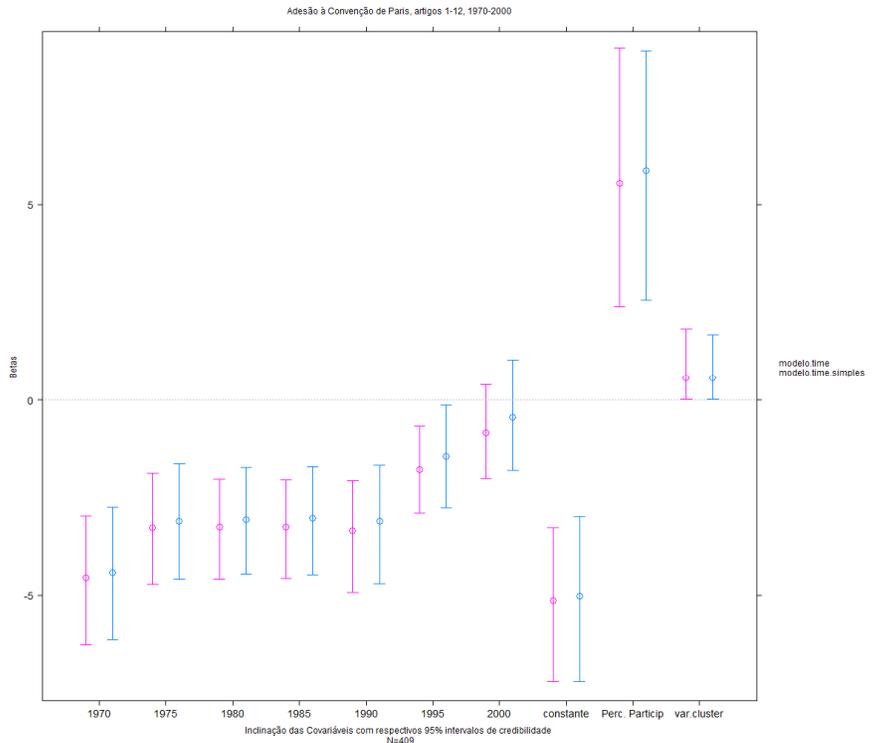
## 5.10. Segundo Nível

Nós voltamos nossa atenção agora para os resultados, em si mesmo importantes, do segundo nível, isto é, da regressão ao nível dos quinquênios. Nós estimamos um modelo em que os interceptos variam por quinquênio, de forma a refletir a tendência temporal de maior adesão dos países ao longo do tempo. Essa modelagem permitiu também incluir uma variável sistêmica apontada pela literatura como relevante, a saber: a taxa de participação de outros estados no tratado. Segundo a literatura (Bernauer, Kalbhen, Koubi e Spilker, 2010; Baccini e Koenig-Archibugi, 2010; Frank, 1999), quanto mais países participam de um tratado, maior a probabilidade dos demais países se engajarem no regime internacional. Uma razão para tanto é que a maior parte dos tratados possui características de bem público, e portanto, tem seus benefícios aumentados e os custos diminuídos à medida que mais e mais países participam do tratado. De fato quanto mais países aplicarem o direito de prioridade e a cláusula de tratamento nacional, maior o benefício para os países participantes do tratado. Assim, essa variável preditora sistêmica foi incluída na estimação do intercepto por quinquênio.

Além disso, um modelo com intercepto variando permite modelar a ideia segundo a qual o fim da Guerra Fria e a passagem de um mundo bipolar para um mundo unipolar alterou a lógica dos regimes internacionais. Kessler (2006) argumenta, por exemplo, que a existência de apenas uma única superpotência pode ter contribuído para a proliferação nuclear e enfraquecimento do regime de não-proliferação nuclear. Skalnes (2006) argumenta que após a guerra fria a ausência de uma grande potência poderia levar a política doméstica a ter mais força sobre fatores sistêmicos nos EUA. Walt (2009) argumenta que potências médias podem procurar formar alianças para contrabalançar o poder do país hegemônico e estados menores tenderiam a se aliar ao país mais hegemônico.

Ao estimar interceptos variando por quinquênio podemos captar esses efeitos sistêmicos sobre os países que operam com o passar do tempo, bem como quaisquer tendências temporais. Porém, sem variáveis que especifiquem esses fatores sistêmicos, não é possível

identificar causalmente quais fatores determinam a tendência temporal observada de mais países aderirem ao longo do tempo. O gráfico 5.10 apresenta os resultados da estimação dos interceptos.



**Gráfico 5.10.** O gráfico apresenta interceptos estimados (variando por quinquênio). Os pontos representam o efeito médio e as barras o intervalo de credibilidade a 95%. A linha pontilhada cruza o zero e facilita a visualização da significância estatística. Barras que não cruzam a linha pontilhadas indicam variáveis significativas. Os modelos 1 e 2 são os definidos anteriormente, no gráfico 5.9. Cada ano representa um intercepto. A variável “Perc.Particip” representa o percentual de estados participando no tratado no quinquênio anterior e foi incluída como preditor do efeito temporal.

Como se pode perceber pelo gráfico 5.10, a variação dos interceptos, medida pela variável “var.cluster”, é pequena, mas positiva. Isso sugere que há algum ganho em adotar um modelo multinível, pois os interceptos diferem, ainda que pouco. De fato, sabemos que quando a variância dos interceptos se aproxima de zero temos essencialmente um modelo *pooled* em que os interceptos não variam (Gelman e Hill, 2007). Se a variância se aproxima de infinito, temos um modelo *unpooled*, em que cada período tem seu próprio intercepto (Gelman e Hill, 2007). Quando estamos no meio termo, o modelo multinível é útil, pois estima os dados levando em consideração as informações no interior de cada ano e também as informações dos outros anos. Pelo gráfico 5.10, vemos que as principais diferenças estão entre os anos 70 e final dos anos 90, onde a probabilidade de adesão é muito maior que no período estudado.

Esses achados refletem a tendência temporal de mais e mais países aderirem aos tratados ao longo do tempo, bem como o aumento na adesão a tratados de propriedade intelectual após o fim da Guerra Fria e a vitória ideológica do Capitalismo sobre o Socialismo.

## 5.11. Ajuste do Modelo

Para além da significância estatística dos coeficientes, é importante olhar para o ajuste do modelo. Medidas de ajuste do modelo em regressão logística são mais complicadas, pois não há uma medida análoga confiável ao  $R^2$  da regressão linear. Além disso, por se tratar de um modelo multinível, medidas como o AIC – *Akaike Information Criterion* - não são pertinentes.

Além do DIC, reportado anteriormente, decidimos utilizar *posterior predictive checks* (Gelman e Hill, 2007). De acordo com Gelman (2003), *model checking* pode ser visto como “the comparison of data to replicated data under the model” (p. 370). E o objetivo da checagem do modelo

*is not the classical goal of identifying whether the model fits or not (and certainly not the goal of classifying models into correct or incorrect, which is the focus of the Neyman-Pearson theory of Type-1 and Type-2 errors), but rather to understand in what ways the fitted model departs from the data (p. 370).*

Para fazer a checagem do modelo, simulamos dados  $y_{.i}^{rep}$  de acordo com o efeito médio estimado pelo modelo e valores amostrais das variáveis para cada entrada país-ano do banco de dados. Para o primeiro modelo, utilizamos:

$$Y_{.i}^{rep} = \text{Bernoulli}(p_{it}),$$

em que:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{renda} * renda_{it} + b_{educ} * educ_{it} + b_{patente} * patente_{it} + b_{legis} * legis_{it})$$

Para o segundo modelo:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{renda} * renda_{it} + b_{democracia} * democracia_{it} + b_{educ} * educ_{it} + b_{patente} * patente_{it} + b_{legis} * legis_{it} + b_{fdi} * fdi_{it})$$

Então, comparamos as previsões dos dados simulados  $y_{.i}^{rep}$  com os dados verdadeiros  $y^{true}$ . Em seguida, computamos a diferença entre as previsões do modelo e os dados reais,  $|y_{.i}^{rep} - y_i^{true}|$ . A taxa de acerto das previsões do modelo é então definida por  $1 - \frac{\sum_{i=1}^n |y_{.i}^{rep} - y_i^{true}|}{n}$ , em que  $n$  é o número de observações do modelo. Como a geração dos dados simulados segue uma Bernoulli, simulamos 100 Bernoullis para levar em conta o caráter estocástico da Bernoulli.

A taxa de acerto média do primeiro modelo foi de 71%. A taxa média de acerto de verdadeiro-positivos, definida pelo total de acertos nas previsões de adesão em relação ao total de adesões, é de 60%, e a taxa média de verdadeiro negativo, definida pelo total de acertos nas previsões de não-adesão em relação ao total de adesões, é de 80%. Esses números indicam que o modelo acerta bastante para prever não adesão, porém erra um pouco mais na previsão de adesão. Isso é de certa forma esperado, na medida em que temos muito menos adesão - 24% dos dados são de adesão.

Os resultados para o segundo modelo foram similares. A taxa de acerto média do segundo modelo foi de 72%, a taxa média de acerto de verdadeiro-positivos é de 57,5%, e a taxa média de verdadeiro negativo, é de 81%. Portanto, o modelo com a inclusão das variáveis não-significativas não apresenta melhoras na previsão intra-amostra.

Se compararmos esses resultados com um modelo ingênuo que tomasse como probabilidade de adesão 24%, ou seja, a proporção de países participantes, a taxa média de acerto seria de 37% e a taxa de verdadeiro-positivo de 25%.

Temos evidência, portanto, de que nosso modelo se ajusta bem aos dados. Não obstante, esse tipo de checagem não leva em conta a incerteza dos parâmetros do modelo. Para tomar em consideração a incerteza do modelo, é preciso realizar simulações não apenas com o efeito médio, mas com os vários valores possíveis dos parâmetros tal como caracterizado pela densidade a posteriori dos mesmos. Felizmente, na inferência Bayesiana estimada por MCMC, nós temos dados que são realizações da densidade a posteriori dos parâmetros. Assim, é possível generalizar a checagem do modelo acima não somente para o efeito médio, mas para todo o intervalo de valores possíveis dos parâmetros caracterizados pela posteriori. Esse procedimento é o que Gelman e Hill (2007) denominam de *posterior predictive checks*. Denotando os dados simulados para a *posterior predictive check* por  $Y_{ij}^{rep}$  (e não mais  $Y_{.i}^{rep}$ ), temos:

$$Y_{ij}^{\text{rep}} = \text{Bernoulli}(p_{itj}),$$

em que:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{\text{renda}} * \text{renda}_{it} + b_{\text{educ}} * \text{educ}_{it} + b_{\text{patente}} * \text{patente}_{it} + b_{\text{legis}} * \text{legis}_{it})$$

O índice  $j$  indica uma realização da densidade a posteriori para o parâmetro<sup>17</sup>.

Se levarmos em conta que os dados são do tipo binário e as estimativas retornam uma probabilidade de adesão, é natural que haja alguns erros apenas por questões probabilísticas. Assim, por exemplo, se o modelo prediz que um dado país num dado ano deve aderir com 90% de probabilidade, e o país de fato aderiu naquele ano, há 10% de chance de erro. Ao reportar um histograma, levamos em conta essa variabilidade.

Os gráficos 5.11a e 5.11b apresentam as taxas de acerto e de verdadeiro-positivo do modelo do *posterior predictive check*. No gráfico 5.11a vemos que a taxa de acerto é maior que 75% e se concentra em torno de pouco mais de 80%. No gráfico 5.11b a taxa de verdadeiro positivo é menor e com variabilidade maior, porém consistentemente maior do que a taxa de acerto do modelo ingênuo, reportada anteriormente.

Assim, a validação intra-amostra apresenta evidências de que o modelo é razoavelmente bem ajustado aos dados. Naturalmente há espaços para melhoria no modelo, especialmente no que tange à previsão de adesão. De todo modo, como a taxa de acerto quanto à adesão não é tão baixa, isso sugere que a decisão de retirar da amostra os dados dos países após à adesão não inviabilizou um modelo preditivo desse evento, menos freqüente.

---

<sup>17</sup> Uma realização de uma variável aleatória é o mesmo que um sorteio (ou no presente caso, simulação) de uma observação da variável aleatória.

Gráfico 5.11a. Posterior Predictive Check  
taxa de acerto - Convenção Paris

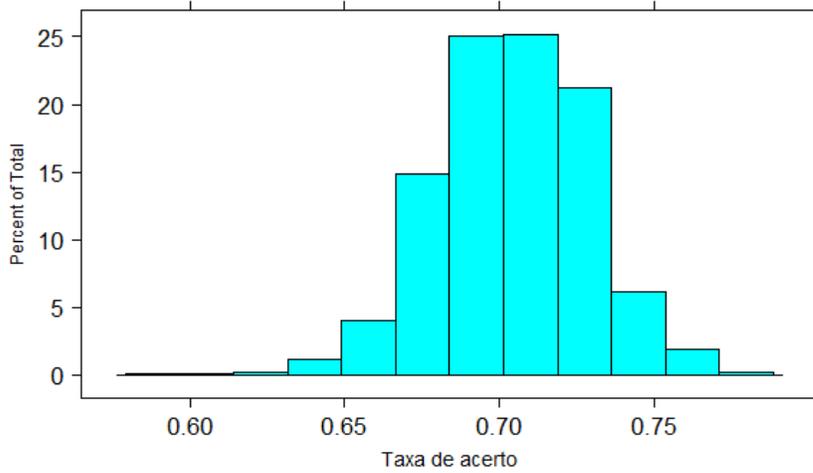


Gráfico 5.11a. O gráfico apresenta o histograma da taxa de acerto do modelo para cada um dos valores amostrados por meio de MCMC para os parâmetros do modelo. Pelo histograma, percebe-se que a taxa de acerto está entre 65% e 75%.

Gráfico 5.11b Posterior Predictive Check  
verdadeiro positivo - Convenção de Paris

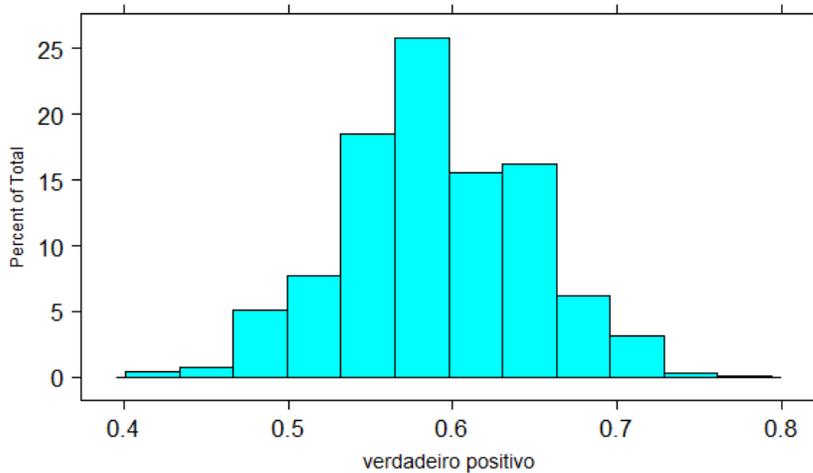


Gráfico 5.11b. O gráfico apresenta o histograma da taxa de verdadeiro positivo do modelo – prevê que o estado aderiu e de fato o estado aderiu - para cada um dos valores amostrados por meio de MCMC para os parâmetros do modelo. Pelo histograma, percebe-se que a taxa de acerto está entre 45% e 75%.

## 6. Resultados Empíricos – Tratado de Cooperação de Patentes

O objetivo deste capítulo é estimar os determinantes da adesão ao Tratado de Cooperação de Patentes e compará-los com os resultados do capítulo anterior, sobre os determinantes da adesão à Convenção de Paris. Nesse sentido, os resultados da análise da adesão ao TCP diferem dos resultados da Convenção de Paris. Nós apresentamos evidências de que os principais determinantes da adesão são os pedidos de patentes no exterior, nível educacional, e renda per capita. Legislação doméstica e IED não são significativos e abertura comercial é significativa ao nível de 90%.

Nós apresentamos fortes evidências de que quanto maior a renda per capita, maior a probabilidade esperada de adesão ao TCP em nossa amostra. Similarmente, quanto maior o logaritmo do número de patentes no exterior, maior a probabilidade de adesão. Nós também encontramos fortes evidências de que para países com renda per capita abaixo de US\$ 6000, quanto maior o nível educacional, menor a probabilidade de adesão. Nossa amostra não permitiu identificar o efeito da educação para países com alto nível de renda per capita, embora haja evidências de que esse efeito possa ser positivo.

Esses achados sugerem que para o TCP, um tratado voltado fundamentalmente para reduzir custos de transação e sem muitos custos para os países que o adotam, são os benefícios diretos ou indiretos que devem governar o comportamento dos países. Países mais desenvolvidos e com mais patentes nos EUA são naturais beneficiários de um acordo como o TCP e, portanto, são esperados os resultados encontrados para essas variáveis.

Nossa interpretação para os resultados do nível educacional e abertura comercial é que países com pouco a ganhar diretamente com a adesão ao tratado pode escolher aderir a ele para tentar obter benefícios indiretos – ou colaterais, na terminologia da literatura. Assim, países com pouca escolaridade têm pouca perspectiva de produzir inovação domesticamente, mas podem desejar aderir ao tratado para sinalizar para a comunidade internacional de que são países que apóiam os regimes internacionais.

Similarmente, países com baixo grau de abertura comercial podem querer ampliar a integração à economia mundial e utilizam a adesão a esse tratado para induzir essa integração.

No entanto, essa interpretação, ainda que de acordo com a literatura, não é baseada em fortes evidências empíricas, de modo que mais estudos devem ser feitos sobre essas variáveis.

A discussão dos resultados está organizada como segue: inicialmente apresentamos estatísticas descritivas para os preditores e possíveis problemas de mensuração; em seguida mostramos os resultados da regressão ao nível dos países e os resultados ao nível dos quinquênios. Posteriormente discutimos o ajuste do modelo. Por fim, comparamos qualitativamente as diferenças de resultado dos determinantes da adesão ao TCP e à Convenção de Paris

## **6.1. Métodos e procedimentos empíricos**

Com vistas a estimar os determinantes da adesão ao TCP, tomamos como variável dependente a data de ratificação do tratado. Nos casos em que essa informação não estava disponível, utilizamos a data de entrada em vigor do tratado<sup>18</sup>. Similarmente ao estudo do caso da Convenção de Paris, o período analisado compreende os anos de 1970 até 2000, com dados quinquenais.

A amostra consiste de 409 observações no formato país-ano, com 90 países desenvolvidos e em desenvolvimento incluídos na análise. Similarmente ao procedimento adotado na análise da adesão à Convenção de Paris, decidimos retirar da amostra os países após eles aderirem ao tratado. As ressalvas a essa abordagem feitas no capítulo anterior para a Convenção de Paris são válidas aqui, *mutatis mutandis*.

As variáveis preditoras ao nível do país incluídas no modelo são as mesmas da Convenção de Paris: i) renda per capita; ii) regime político doméstico; iii) nível educacional do país; iv) patentes nos EUA; v) grau de proteção da legislação doméstica de patentes; vi) fluxo de investimento externo direto como proporção do PIB; vii) abertura comercial.

Abaixo temos algumas estatísticas descritivas de cada variável preditora, além detalhes do processo de mensuração das mesmas.

## **6.2. Renda per capita**

A medida para renda per capita é a mesma utilizada na análise da Convenção de Paris. Nós padronizamos a variável durante o processo de estimação, ao subtrair a média de todos os países em todo o período e dividir pelo desvio padrão. A padronização da variável permitiu

---

<sup>18</sup> Ver o capítulo 3, em que discutimos em mais detalhes a operacionalização empírica da variável dependente.

que a convergência das simulações se desse de forma mais eficiente, de forma que essa especificação foi empregada em todas as simulações.

Na nossa amostra, a renda per capita média é de US\$ 7058 e a mediana é US\$ 5225 e o desvio-padrão é de US\$ 6018. O gráfico 6.2 apresenta um box-plot da renda para os países que aderiram e não aderiram ao TCP:

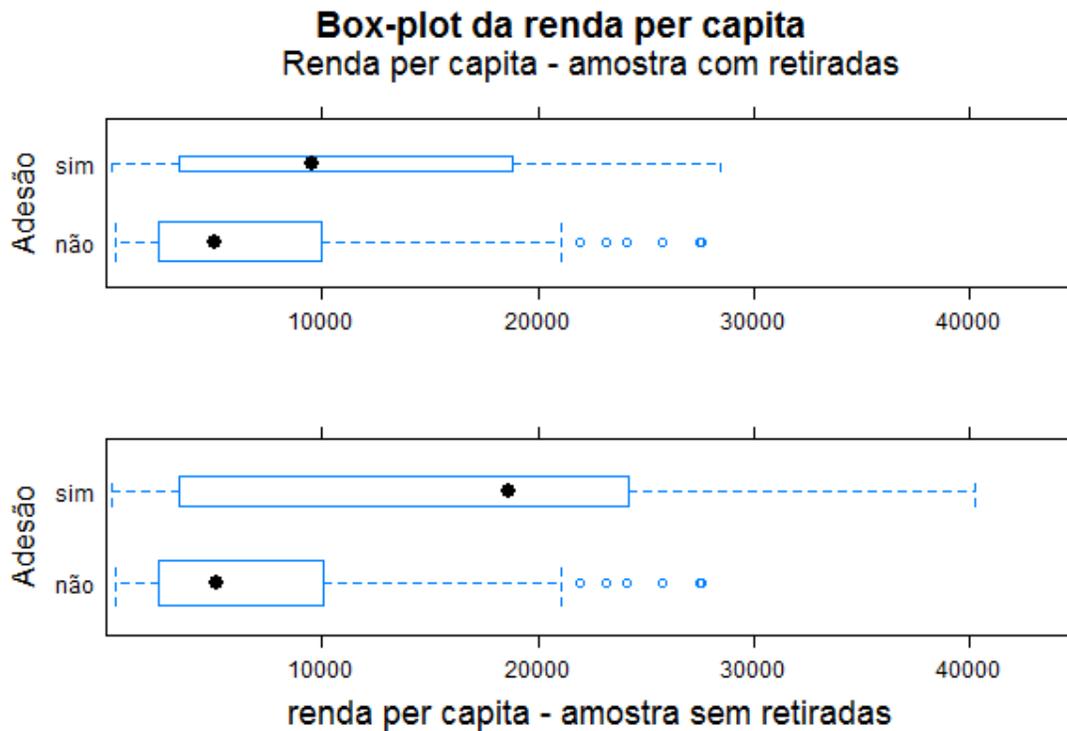


Gráfico 6.2. Box-plot renda per capita dos países por adesão e não adesão ao TCP. O painel superior apresenta os dados para a amostra com retiradas e o painel inferior para a amostra sem retiradas. No painel superior, percebe-se uma associação positiva entre adesão e renda per capita. Comparando com o painel inferior, percebe-se que essa associação é subestimada.

O gráfico 6.2, painel superior, mostra a distribuição da renda segundo a adesão ou não dos países. Como se vê, a renda mediana dos países que aderiram é maior que a dos que não aderiram. Similarmente, o terceiro quartil dos países participantes do tratado é maior do que entre não participantes. Para checar a possibilidade de viés de seleção, a comparação entre a amostra com retiradas (painel superior) e sem retiradas (painel inferior) mostra que o efeito é subestimado.

Na amostra sem retirada (painel inferior do gráfico 6.2), a mediana amostral dos estados que aderem ao TCP é maior do que a mediana amostral dos aderentes ao tratado na amostra com retiradas (painel superior). É possível, portanto, que haja algum viés na estimação da

associação entre renda e probabilidade de adesão. Porém, como esperávamos teoricamente que a associação fosse positiva e as estatísticas descritivas apontam nessa direção, é possível afirmar que o viés deve subestimar o efeito da associação positiva. Podemos então identificar limites inferiores para o efeito da renda sobre a probabilidade de adesão na análise regressão.

### 6.3. Regime Político

A variável “regime político” é a mesma utilizada anteriormente e foi operacionalizada da mesma maneira. Como a mostra é um pouco diferente, reportamos aqui também um *mosaic plot* para visualizar a relação entre democracia e adesão.

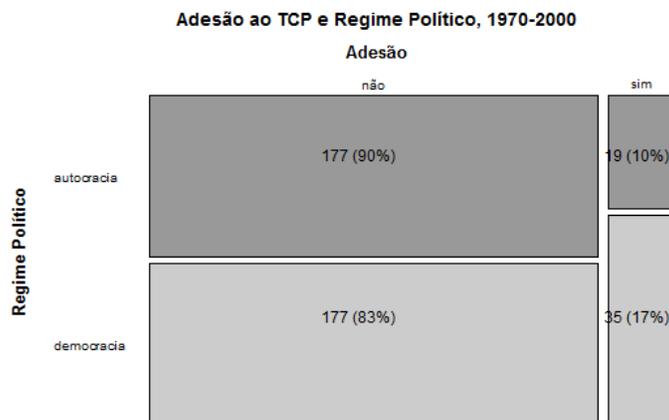


Gráfico 6.3. Mosaic Plot de regime político e adesão ao TCP. O Mosaic plot é uma representação gráfica das proporções de adesão por Regime Político. Assim, do total de autocracias da amostra, 90% não aderiram ao TCP, enquanto 10% aderiram. E do total de democracias, 83% aderiram, e 17% aderiram. Essa maior proporção de democracias que aderem ao TCP em relação a autocracias é visualizada no gráfico pelo fato da área do retângulo representando democracias que aderem (“sim”) ser maior do que a área de autocracias que aderem. Por outro lado, a área de autocracias que não aderem (“não”) é maior do que a área de democracias que não aderem.

O gráfico 6.3 apresenta um **mosaic plot**, que é uma representação gráfica da tabela de contingência. O gráfico mostra se regime político está associado à maior ou menor adesão. Os números apresentam o total de observações e entre parênteses o percentual de observações de cada regime. Assim, vemos que 90% das autocracias não aderem, e 10% aderem ao tratado. Entre países democráticos, 83% não aderem ao tratado, enquanto apenas 17% aderem. Visualmente é fácil perceber que temos mais democracias aderindo do que autocracias, enquanto o inverso é verdadeiro para não adesão: mais autocracias aderem do

que democracias. De fato, essa diferença é significativa ao nível de 5%. Assim, diferentemente da Convenção de Paris, democracias aderem mais que autocracias ao TCP.

## 6.4. Nível Educacional

Para mensurar o nível educacional, utilizamos os dados disponíveis em Barro e Lee (2000), tal como na análise da Convenção de Paris.

O gráfico 6.4 apresenta a distribuição do nível de educação entre países que aderiram e não aderiram ao TCP. É fácil ver que os países que aderem ao tratado têm maior nível de escolaridade, embora haja vários *outliers* com alta escolaridade que não aderem ao TCP. Numa regressão simples é possível constatar que essa diferença é significativa, indicando que países com maior nível de escolaridade têm maior probabilidade de adesão ao TCP.

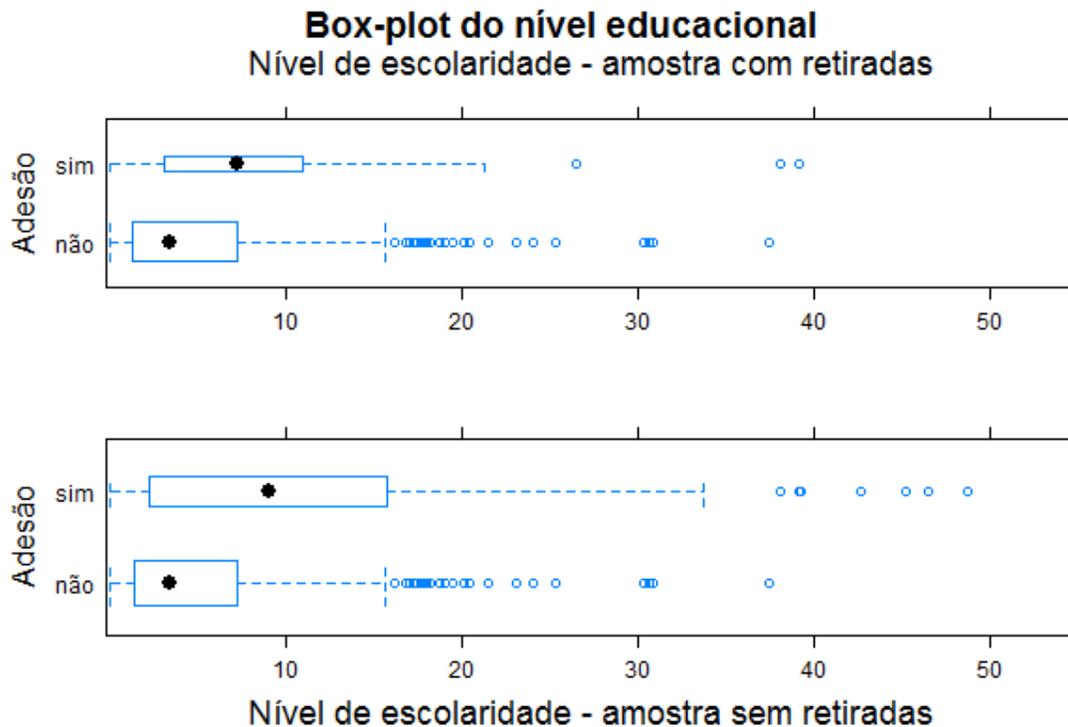


Gráfico 6.4. Box-plot do nível de escolaridade dos países por adesão e não Adesão ao TCP. O painel superior apresenta os dados para a amostra com retiradas e o painel inferior para a amostra sem retiradas. No painel superior, percebe-se uma associação positiva entre adesão e educação. Comparando com o painel inferior, percebe-se que essa associação é subestimada.

Os gráficos 6.4a e 6.4b apresentam os dados de educação por quinquênio na amostra com retirada de países e sem retiradas (amostra completa):

Devido ao baixo número de observações de adesão para os primeiros anos, a visualização gráfica é um pouco difícil para os anos 70. Para os anos seguintes, não é possível perceber diferenças significativas entre países que aderem e que não aderem em ambos os gráficos. Isso indica que a retirada de países da amostra não está produzindo um desbalanceamento da amostra.

## **6.5. Legislação doméstica em patentes**

Conforme argumentado no segundo capítulo, utilizamos o índice Ginarte-Park, consagrado pela literatura, para mensuração do grau de proteção patentária dos países. Na análise estatística decidimos utilizar uma variável defasada em um quinquênio, pois esperamos que mudanças em períodos anteriores tivessem seus impactos apenas nos anos seguintes, e não no próprio ano. Os detalhes das modificações empreendidas por nós para tornar os dados adequados aos nossos propósitos encontram-se no referido capítulo. Vale destacar que nós *imputamos* os valores do Reino Unido para o período pré-1995, pois não tínhamos os dados desagregados para o período 1970-1990. Porém, como Park (2008) provê o valor médio do índice para esse período, decidimos repetir esse valor médio para os anos 1965-1990. Embora haja imprecisão nesse método, no período considerado a maior parte dos países desenvolvidos não alterou sua legislação doméstica de patentes, de forma que o índice para esses países é quase constante nesse período. Assim, não é uma aproximação tão ruim quanto pode parecer à primeira vista.

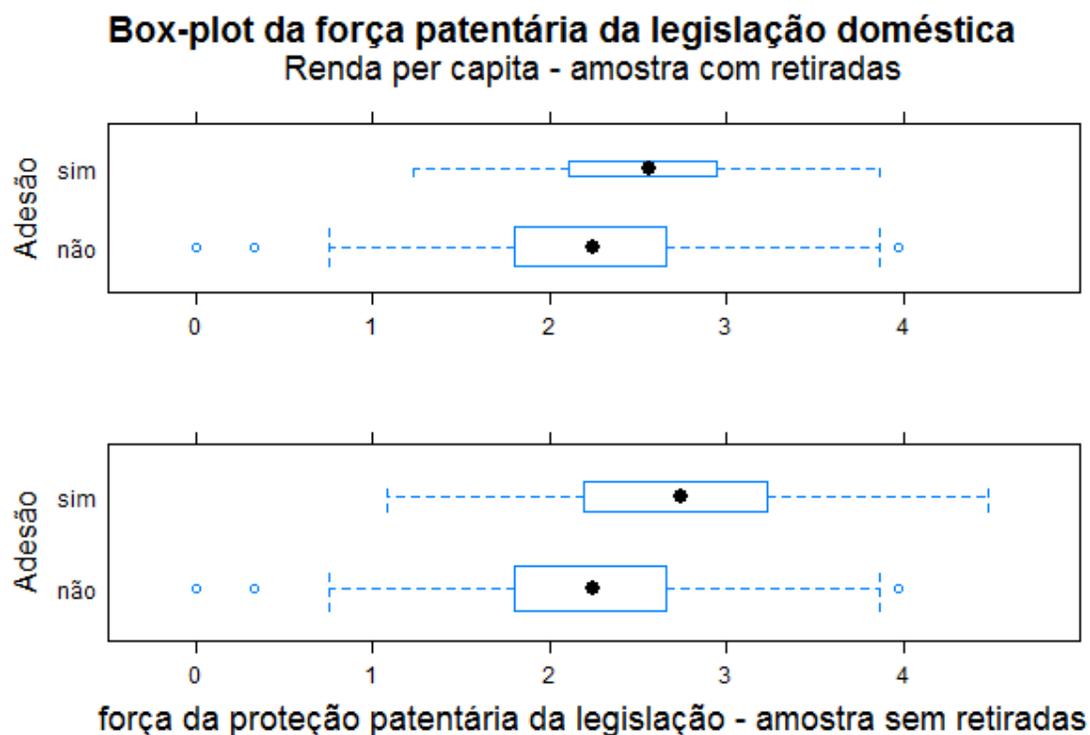


Gráfico 6.5. Box-plot da força de proteção patentária da legislação doméstica, medida pelo índice Ginarte-Park, dos países por adesão e não Adesão ao TCP. O painel superior apresenta os dados para a amostra com retiradas e o painel inferior para a amostra sem retiradas. No painel superior, percebe-se uma associação positiva entre adesão e força da proteção patentária doméstica. Comparando com o painel inferior, percebe-se que essa associação é pouco subestimada.

O gráfico 6.5 apresenta a distribuição dos dados do índice Ginarte-Park para todo o período, e por ano na amostra com retiradas e na amostra completa. Como se pode observar no gráfico 6.5, os dados são bem distribuídos e com poucos *outliers*. Países com legislação patentária mais forte tendem a aderir mais ao TCP do que países com legislação mais fraca.

Analisando os gráficos, não parece haver grandes diferenças entre os dados. Assim, não nos parece que a retirada de países produza vieses significativos, embora se houver é um viés negativo, de subestimação do efeito.

## 6.6. Patentes nos Estados Unidos

Como *proxy* para a importância de obter proteção patentária no estrangeiro, utilizamos o número de patentes obtidas nos EUA. Os dados foram retirados do website do US Patente Trade Office (USPTO, 2009). Para cálculo dos dados quinquenais, utilizamos a média de cinco anos, similarmente à renda per capita.

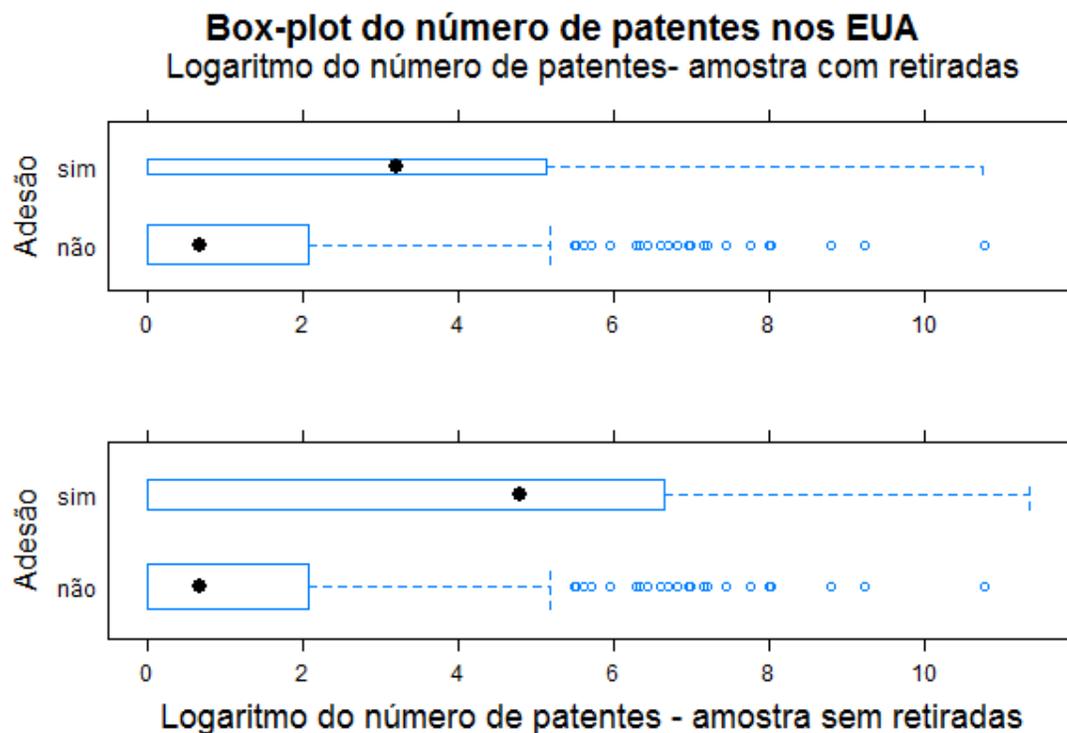


Gráfico 6.6. Box-plot do logaritmo do número de patentes nos EUA, dos países por adesão e não adesão ao TCP. O painel superior apresenta os dados para a amostra com retiradas e o painel inferior para a amostra sem retiradas. No painel superior, percebe-se uma associação positiva entre adesão e patentes nos EUA. Comparando com o painel inferior, percebe-se que essa associação é subestimada.

O gráfico 6.6 apresenta a distribuição dos dados agregados do logaritmo do número de patentes nos EUA. A mediana dos países que aderiram é maior do que a dos que não aderiram. Além disso, as patentes obtidas no exterior pelos países que aderem são mais dispersas no extremo superior da distribuição do que entre os que não aderem.

Analisando os mesmos dados por ano na amostra sem retiradas de países (painel inferior) e com retiradas (painel superior), vemos que a mediana dos aderentes é um pouco reduzida. O intervalo interquartil é diminuído, conforme o esperado. Assim, é possível que haja algum viés subestimando os dados. As evidências apontam que pelo menos os limites inferiores na regressão são identificáveis.

## 6.7. Investimento Externo Direto (IED)

Os dados de investimento externo direto (IED) foram retirados do bando de dados do Banco Mundial (2010). Os dados de investimento externo direto refletem o fluxo líquido de investimentos adquiridos pelo menos 10% do capital votante de empresas operando em outro país que não o do investidor Banco Mundial (2010). Pode ser pensado também como a soma

do *equity capital*, reinvestimento de lucros e outros capitais de longo e curto prazo tal como mostrado nos Balanços de Pagamento dos Países. Os dados são em dólares correntes. Assim, dados de IED podem tanto ser positivos como negativos. IED negativo significa que houve mais desinvestimento externo do que investimento externo.

Como o tamanho da economia de um país afeta o montante de IED que ela pode receber, parece-nos mais conveniente utilizar a proporção de IED em relação ao PIB como medida do fluxo de investimento externo direto. Para obter os dados de cada quinquênio, nós dividimos a média dos cinco anos do fluxo de investimento externo direto pela média do PIB do país no mesmo período, medido em poder de paridade de Compra.

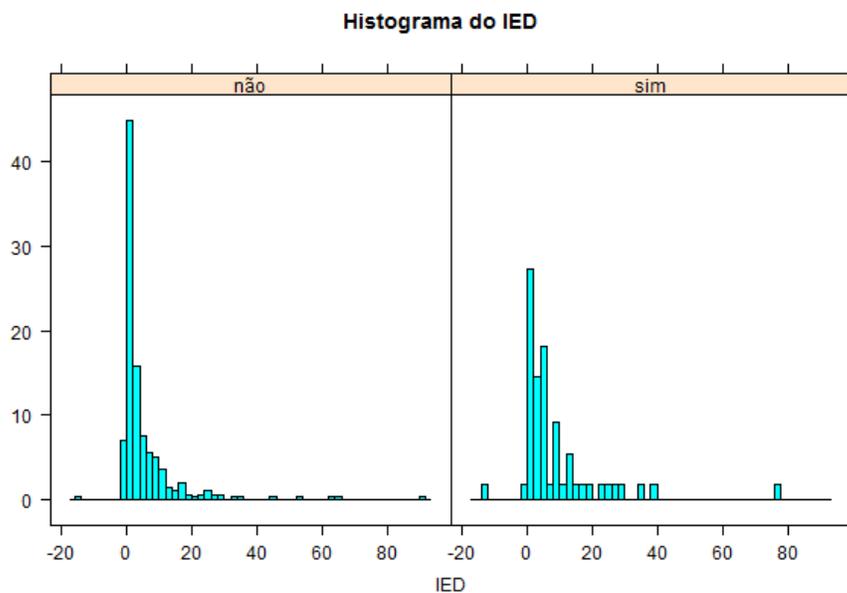


Gráfico 6.7. Histograma para o Investimento Externo Direto em relação ao PIB entre países que aderiram (sim) e não aderiram (não) ao TCP, na amostra com retiradas. Vemos que praticamente há uma pequena diferença na distribuição entre estados aderentes e não aderentes ao tratado. O IED é levemente maior entre estados membros do TCP que entre não membros. O valor modal é próximo do zero, indicando baixo fluxo de investimento externo direto.

O gráfico 6.7 mostra o histograma para o percentual de investimento externo direto em relação ao PIB. A cauda superior do histograma dos países que aderem é um pouco maior do que a dos países que não aderem. Assim, temos evidências de que quanto maior o nível de IED, maior a probabilidade de adesão.

## 6.8. Abertura Comercial

O TCP é um tratado que visa a redução de custos de transação na obtenção de patentes no exterior. Portanto, é um instrumento especialmente útil para empresas com atuações em

vários mercados. Assim, países com economia mais integrada à economia mundial podem se beneficiar mais com a participação no TCP, de forma que alguma medida de integração comercial deve ser incorporada no modelo.

Nós incluímos uma medida de abertura comercial, que é a soma das exportações e importações relativa ao PIB do país. Essa variável consta da base *Pen World Table 6.3*, denominada “openk”. Porém, países menores tendem a comercializar mais com o mundo, pois não possuem diversidade industrial doméstica suficiente. Assim, para controlar para o tamanho do país, dividimos a abertura comercial pela população do país. Para usar dados quinquenais, tomamos a média de cinco anos, como nas demais variáveis em que computamos a média quinquenal.

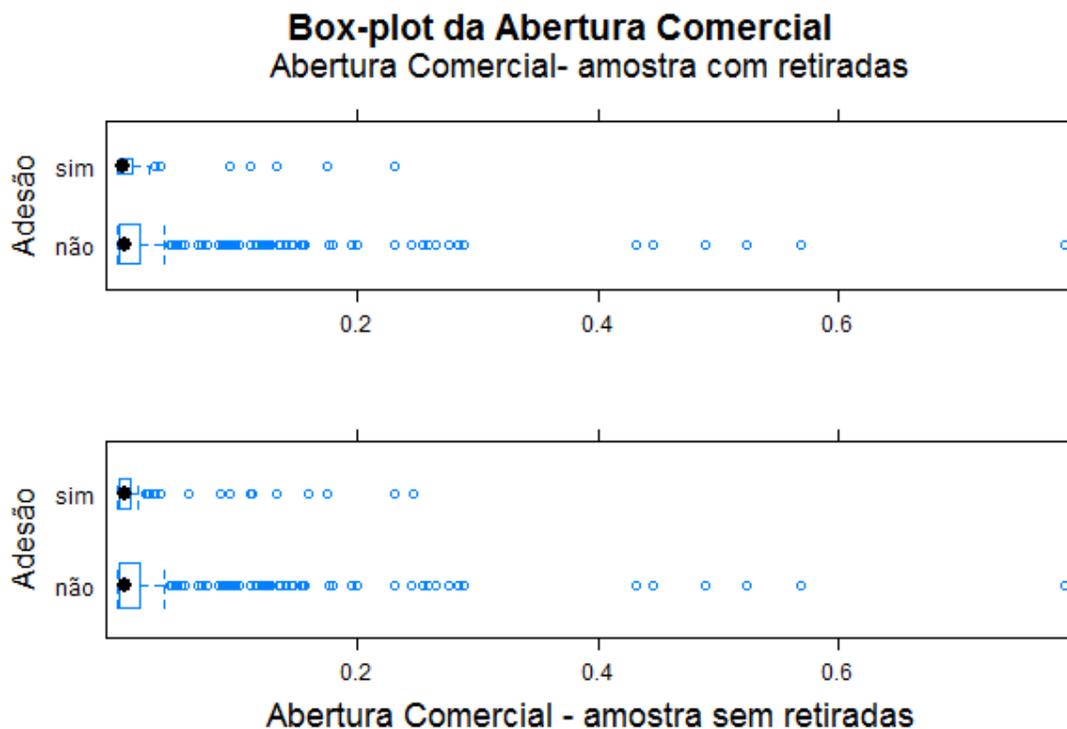


Gráfico 6.8. Blox-plot da abertura comercial (exportações + importações dividida pelo PIB) relativa ao tamanho do país (medida pelo tamanho da população), dos países por adesão e não adesão ao TCP. O painel superior apresenta os dados para a amostra com retiradas e o painel inferior para a amostra sem retiradas. No painel superior, percebe-se não há associação entre adesão e abertura comercial. Comparando com o painel inferior, percebe-se pouca diferença.

O gráfico 6.8 apresenta o Box-plot para abertura comercial. Nele podemos ver que não temos evidência de associação entre abertura e adesão, nem tampouco de vieses introduzidos por nosso procedimento de retirar da amostra países após adesão ao TCP.

## 6.9. Variáveis ao nível do quinquênio

Como estamos estimando um modelo hierárquico de dois níveis com intercepto variando por quinquênio, incluímos como preditor dos interceptos o percentual de países participantes da Convenção de Paris por quinquênio. Para tanto, calculamos, pra cada ano, o percentual de países que aderiram ao tratado em relação ao total de países que poderiam ter aderido. Nessa variável foram incluídos mais do que os 90 países do primeiro nível, pois tínhamos informações para todos os países da base de dados se havia aderido ou não ao tratado. O número total de países, porém, varia no tempo, já que nesse período há tanto a criação como dissolução de estados independentes.

O gráfico 6.9 apresenta a evolução da participação na no TCP. Como se vê, há uma tendência crescente no tempo de adesão. Na estimação do modelo, nós utilizamos a variável defasada em cinco anos, de forma evitar problema de causalidade reversa, já que a adesão num dado ano aumenta a participação daquele ano, criando uma correlação espúria.

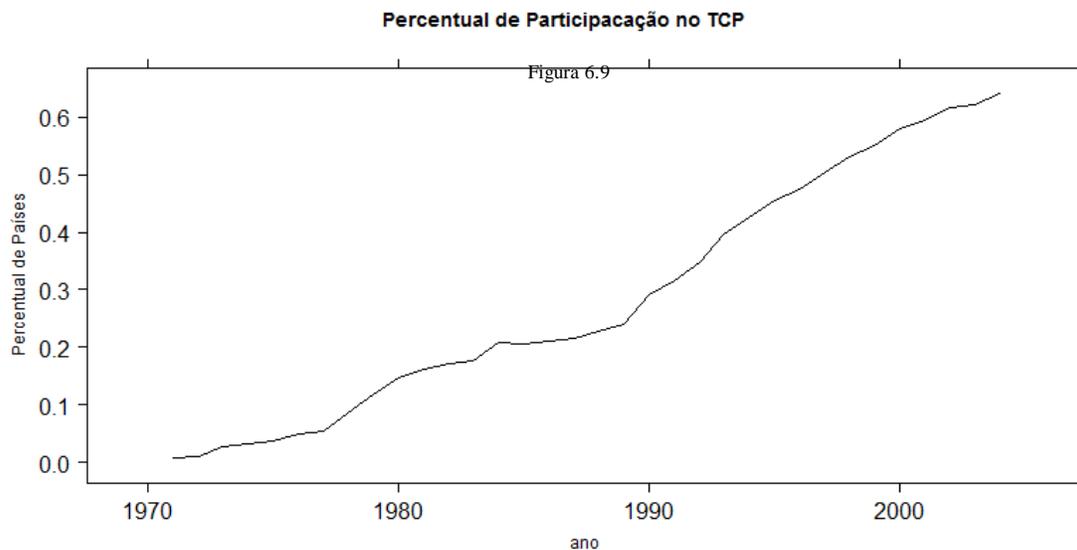
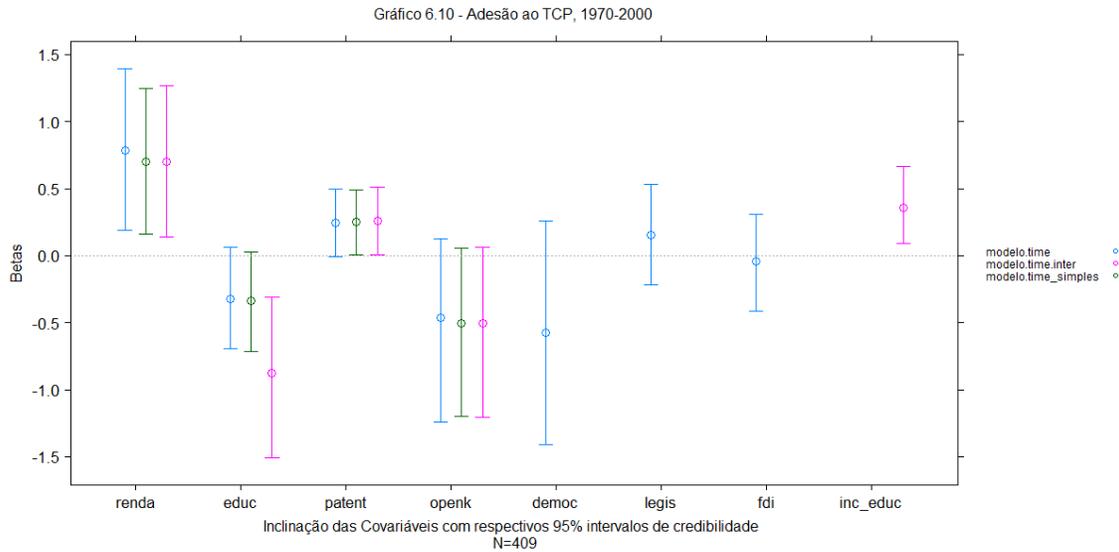


Gráfico 6.9. Percentual de estados membros do TCP por ano. No começo dos anos 70, menos de 1% dos países considerados participavam do tratado. Ao final dos anos 2000, pouco mais de 65% dos estados da amostra haviam aderido ao acordo.

## 6.10. Modelo de Regressão

Nós estimamos uma regressão logística com intercepto variando por ano - modelo hierárquico. Um modelo com intercepto variando por país também foi testado, porém o modelo com intercepto variando por ano teve um ajuste superior. O anexo do trabalho

apresenta alguns resultados do modelo com intercepto variando por país. O gráfico 6.10 abaixo apresenta os resultados nas variáveis independentes do modelo estimado (os interceptos variando por país são mostrados na próxima seção).



**Gráfico 6.10.** O gráfico apresenta os resultados de três regressões Bayesianas hierárquicas no primeiro nível (nível dos países). Os pontos representam o efeito médio e as barras o intervalo de credibilidade a 95%. A linha pontilhada cruza o zero e facilita a visualização da significância estatística. Barras que não cruzam a linha pontilhadas indicam variáveis significativas. As variáveis encontram-se no eixo “x”. Se uma variável não foi incluída no modelo, não há barras da cor do modelo para aquela variável.

O gráfico 6.10 acima mostra os resultados da regressão multinível de três modelos. O primeiro modelo, em azul, é o resultado da regressão com intercepto variando por quinquênio e com todas as variáveis testadas por nós:

$$\Pr(y_{it} = 1) = \text{logit}^{-1}(\alpha_t + b_{renda} * renda_{it} + b_{educ} * educ_{it} + b_{patente} * patente_{it} + b_{openk} * openk_{it} + b_{democracy} * democracy_{it} + b_{legis} * legis_{it} + b_{fdi} * fdi_{it})$$

$$\alpha_t \sim \text{Normal}(g_0 + g_1 * \text{perc. particip}, \sigma_{\alpha}^2)$$

Priors:

$$b_{renda} \sim N(0, 1000); b_{educ} \sim N(0, 1000); b_{patente} \sim N(0, 1000); b_{openk} \sim N(0, 1000);$$

$$b_{democracy} \sim N(0, 1000); b_{fdi} \sim N(0, 1000); b_{legis} \sim \text{Normal}(0, 1000); g_0 \sim N(0, 1000);$$

$$g_1 \sim N(0, 1000); \sigma_{\alpha}^2 \sim \text{Uniforme}(0, 100)$$

O segundo modelo, em verde, é similar ao primeiro, porém sem as variáveis não-significativas do primeiro modelo e com *prioris* iguais para os coeficientes das variáveis:

$$\begin{aligned} \Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha_t + b_{\text{renda}} * \text{renda}_{it} + b_{\text{educ}} * \text{educ}_{it} + b_{\text{patente}} * \text{patente}_{it} \\ + b_{\text{openk}} * \text{openk}_{it}) \\ \alpha_t \sim \text{Normal} (g_0 + g_1 * \text{perc. particip}, \sigma_{\text{alpha}}^2) \end{aligned}$$

O terceiro modelo, em rosa, apresenta um termo de interação entre educação e renda e com *prioris* iguais para os coeficientes das variáveis:

$$\begin{aligned} \Pr (y_{it} = 1) = \text{logit}^{-1} (\alpha_t + b_{\text{renda}} * \text{renda}_{it} + b_{\text{educ}} * \text{educ}_{it} + b_{\text{patente}} * \text{patente}_{it} \\ + b_{\text{openk}} * \text{openk}_{it} + b_{\text{renda.educ}} * \text{renda}_{it} * \text{educ}_{it}) \end{aligned}$$

As barras representam 95% de intervalos de credibilidade e os círculos, a média da posteriori. Assim, variáveis são significativas se as barras não cruzam a linha pontilhada.

No modelo sem termo de interação, vemos que renda per capita e patentes no exterior são as únicas variáveis significativas ao nível de 95% de certeza. Legislação doméstica, educação, regime político, abertura comercial e investimento externo direto não foram significativos.

Tendo em vista a análise descritiva, o modelo surpreende por regime político e nível de escolaridade não serem significativos e também terem o sinal errado. Na análise descritiva, tanto educação quanto democracia foram associadas positivamente com probabilidade de adesão. Após análises, descobrimos que a inclusão da renda per capita torna a associação entre educação e adesão negativa. Como as variáveis têm alta correlação na amostra (60%), decidimos incorporar um termo de interação entre renda per capita e nível educacional em um modelo para tentar entender se podemos explicar o resultado contrário ao esperado.

O ajuste do primeiro modelo, medido pelo DIC, foi de 272,1. O segundo modelo teve um DIC de 263,7, indicando que as demais variáveis efetivamente pouco adicionavam ao modelo. O terceiro modelo, com interação, teve DIC de 264,3, indicando ajuste similar ao segundo modelo. Assim, por essa medida de ajuste do modelo, parece-nos justificada a incorporação do termo de interação. Como veremos os resultados do modelo com interação fazem mais sentido teoricamente e, portanto, é o nosso modelo preferido.

Contudo, termos de interação são difíceis de interpretar visualizando apenas os coeficientes. Essa dificuldade é ainda maior em se tratando de regressão logística, visto que há uma

relação não-linear entre os preditores e a probabilidade de adesão. Desse modo, interpretar diretamente o coeficiente estimado, à maneira da regressão linear, não é aconselhável. Porém, é possível estimar, similarmente à regressão linear, algo como um efeito médio preditivo das variáveis. Gelman e Hill (2007) sugerem computar a diferença preditiva média, que definimos em seguida. Contudo, primeiramente precisamos distinguir conceitualmente *inputs* e preditores. Essa distinção é importante em modelos com interação, mas não em modelos sem interação.

Preditores são todas as variáveis incluídas no modelo, incluindo interações. Assim, num modelo com renda, educação e interação de educação com renda, temos quatro preditores (intercepto, renda, educação e interação renda-educação). Porém, como um dos preditores é a combinação de outras duas variáveis, diremos que temos apenas dois *inputs*, renda e educação. Assim, *inputs* são as variáveis do modelo sem interações e sem contar intercepto.

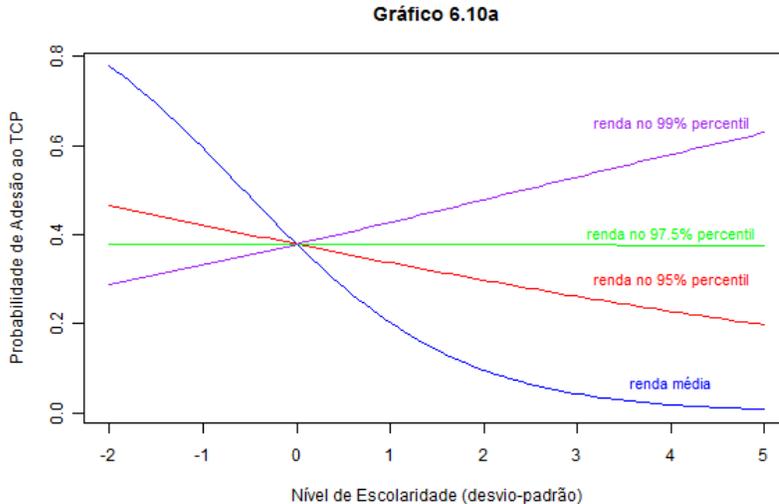
A diferença preditiva é então definida como a diferença esperada na probabilidade de adesão associada com uma mudança no input de interesse, mantendo constantes as demais variáveis (Gelman e Hill, 2007). E a diferença preditiva média é computada tomando a média para todos os valores das variáveis na base de dados. No anexo definimos matematicamente o cômputo da diferença preditiva média. Para nós, o importante é que a diferença preditiva média pode ser interpretada similarmente como o coeficiente estimado da regressão linear.

A diferença preditiva média da renda em 2000, ano com maior efeito, é de 30%. Assim, em média nos nossos dados, no ano 2000, países com renda per capita de US\$ 13.605 são 15% mais prováveis de aderirem ao TCP do que países com renda per capita de US\$ 6.152,00.

A mesma análise pode ser feita para estimar o efeito esperado de variações no nível de escolaridade. A diferença preditiva de um aumento em 2 desvios-padrão na educação no ano 2000 é de -20%. Este resultado implica que nos dados, em média, estados com 18% da população acima de 15 anos com nível maior que secundário são 20% menos prováveis de aderir ao TCP do que estados com 6% de população com nível maior que secundário.

Como o modelo contém um termo de interação, podemos estimar o efeito preditivo da educação sobre probabilidades de adesão para vários níveis de renda. O gráfico 6.10a apresenta o efeito preditivo da educação (em desvios-padrão da média) para países com renda média, renda no 95% percentil, 97,5% percentil e 99% percentil e assumindo valores médios para as demais variáveis do modelo. Para países com renda per capita abaixo do 97,5%

percentil da distribuição, o efeito preditivo médio da educação é negativo. Isso implica que para países com renda per capita menor do que US\$ 22.000 o efeito preditivo esperado da educação é negativo.



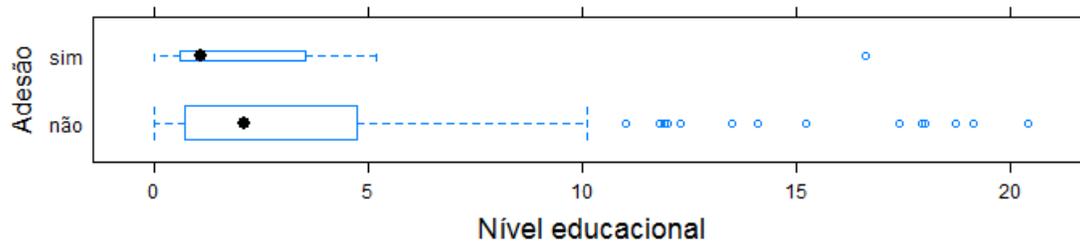
**Gráfico 6.10a. Efeito da escolaridade para vários níveis de renda. O gráfico mostra que, quanto menor a renda, mais o efeito da educação é negativo.**

Esse é, contudo, o efeito médio. Se acrescentarmos intervalos de credibilidade para os coeficientes estimados ao nível de 95% de certeza, então é possível que o efeito preditivo da educação seja negativo para os países no extremo superior da distribuição, bem como é possível que os efeitos negativos da educação sobre adesão se dêem apenas para países com renda per capita abaixo da média.

Se retornarmos aos dados brutos, apresentados no gráfico 6.10.1, veremos que dividindo a amostra entre países com renda per capita menor do que US\$ 6.000,00 e maior do que US\$ 6.000,00 temos a interação predita: para países mais pobres, a adesão é maior entre aqueles com menor nível educacional. Para países mais ricos, a adesão é maior entre aqueles com maior nível educacional.

Por outro lado, notamos anteriormente que possivelmente há viés no efeito de educação. Esse viés permite-nos identificar os limites inferiores, mas não os limites superiores. Em outras palavras, não é possível determinar se o efeito de educação é menor do que zero, pois podemos estar subestimando o valor do coeficiente.

### Box-plot do nível educacional - países com renda menor que US\$ 6.000



### Box-plot do nível educacional - países com renda maior que US\$ 6.000

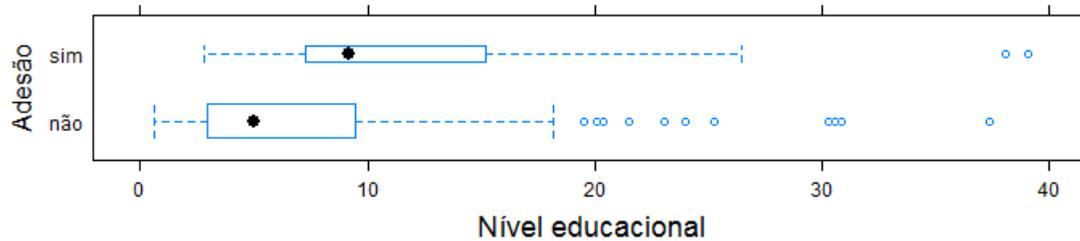
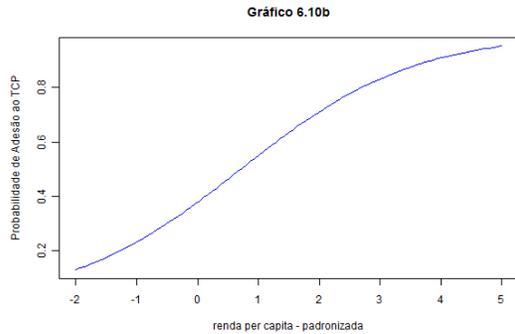


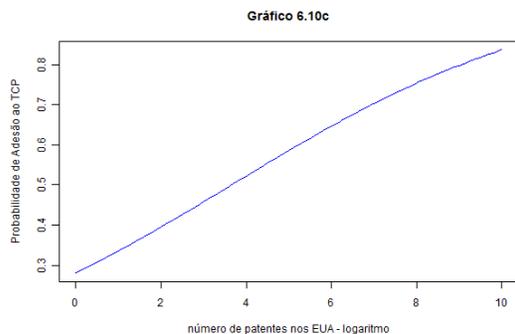
Gráfico 6.10.1. Box-plot do nível educacional. O painel superior apresenta o Box-plot para países com renda per capita abaixo de US\$ 6000,00, e o painel inferior para países com renda per capita maior do que US\$ 6000,00.

Em resumo, de um lado a diferença preditiva média e dados brutos apontam para um efeito negativo, mediado pelo nível de renda dos estados. De outro lado, a possibilidade de viés e as incertezas estimadas sobre o valor dos coeficientes sugerem cautela na interpretação dos resultados. Tudo bem pesado, temos evidências do efeito negativo para países com menor nível de renda e efeito positivo para níveis de renda maiores, porém as evidências não são muito fortes.

Com relação às demais variáveis, como apontamos inicialmente, a renda per capita e o número de patentes nos EUA estão associados positivamente com a probabilidade de adesão. Os gráficos 6.10b e 6.10c apresentam para o ano 2000 o efeito médio da renda per capita e do logaritmo do número de patentes nos EUA, respectivamente.



**6.10.b. Efeito da renda per capita, tudo o mais constante no valor médio. As unidades de medida são em desvio-padrão. Assim, o zero representa a renda per capita média e cada unidade um desvio padrão a mais ou a menos. Vemos que países com renda per capita 2 desvios-padrão abaixo da média têm probabilidade perto de zero de adesão, e países com 4 desvios-padrão acima da média têm probabilidade quase 100%.**



**6.10.c. Efeito do logaritmo do número de patentes nos EUA, tudo o mais constante no valor médio. Para ter uma idéia mais precisa da magnitude, é preciso exponenciar os dados. Assim, zero significa zero patentes, 4 no gráfico representa 55, 6 equivale a 403 patentes, 8 vale 2981 e 10 equivale a 22026 patentes. Vemos que países com zero patentes nos EUA têm probabilidade perto de zero de adesão, e países com 22 mil patentes nos EUA têm probabilidade quase 100%.**

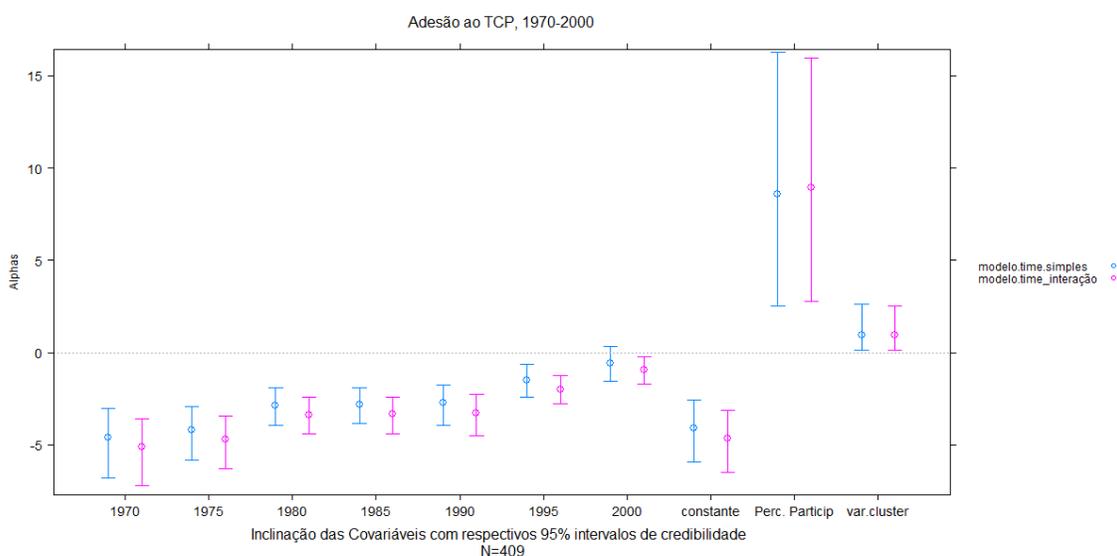
O intervalo de credibilidade para abertura comercial cruza o zero, de forma que não é um resultado significativo ao nível de 95%. Porém, podemos computar a probabilidade do parâmetro ser menor do que zero, como fizemos na análise da Convenção de Paris. De fato, a probabilidade de o parâmetro associado à abertura comercial ser menor do que zero é 96%.

Esse resultado é contra-intuitivo, já que países mais abertos economicamente tenderiam a se beneficiar mais da adesão ao TCP. Como, porém, o resultado é apenas marginalmente significativo e na análise da estatística descritiva não vimos evidências fortes de associação, temos apenas evidências fracas de que países mais abertos comercialmente têm menor probabilidade de adesão ao TCP.

## 6.11. Segundo Nível

Nós voltamos nossa atenção agora para os resultados, em si mesmo importantes, do segundo nível, isto é, da regressão ao nível dos quinquênios. Nós estimamos um modelo em que os interceptos variam por quinquênio, de forma a refletir a tendência temporal de maior adesão dos países ao longo do tempo. A *rationale* para tal é similar a apresentada para o modelo estimado para a Convenção de Paris.

Assim, ao estimar interceptos variando por quinquênio podemos captar efeitos sistêmicos sobre os países que operam com o passar do tempo, bem como quaisquer tendências temporais. O gráfico 6.11 apresenta os resultados da estimação dos interceptos.



**Gráfico 6.11.** O gráfico apresenta interceptos estimados (variando por quinquênio). Os pontos representam o efeito médio e as barras o intervalo de credibilidade a 95%. A linha pontilhada cruza o zero e facilita a visualização da significância estatística. Barras que não cruzam a linha pontilhada indicam variáveis significativas. Os modelos 1 e 2 são os definidos anteriormente, no gráfico 6.10. Cada ano representa um intercepto. A variável “Perc.Particip” representa o percentual de estados participando no tratado no quinquênio anterior e foi incluída como preditor do efeito temporal.

Como se pode perceber pelo gráfico 6.11, a variação dos interceptos, medida pela variável “var.cluster” é positiva e significativamente maior do que zero. Isso sugere que há algum ganho em adotar um modelo multinível, pois os interceptos diferem, ainda que pouco. Pelo gráfico 6.11, vemos basicamente três grupos de tendências temporais: anos 70 e 75; anos 80, 85 e 90, e anos 95 e 2000.

Grosso modo, esses períodos correspondem a: i) fase inicial do tratado em que ele ainda não havia entrado em vigor e, portanto, o incentivo à entrada de outros países era baixo; ii) anos 80, 85 e 90, ainda durante a Guerra Fria, pré-OMC e independência das ex-repúblicas

soviéticas; e iii) 95 e 90, pós-OMC e surgimento de estados no Leste Europeu e com elevada participação internacional no tratado.

Esses achados refletem a tendência temporal de mais e mais países aderirem aos tratados ao longo do tempo. Ademais, esses achados sugerem efetivamente que temos efetivamente efeitos sistêmicos diferentes ao longo do tempo, possivelmente resultado das transformações no sistema internacional, particularmente o fim da Guerra Fria e a criação da OMC e TRIPs, que são reconhecidamente importantes para o regime internacional de patentes.

## 6.12. Ajuste do Modelo

Para analisar o ajuste do modelo, utilizamos o mesmo procedimento do capítulo anterior (Seção 5.11).

Para fazer a checagem do modelo, simulamos dados  $y_{it}^{rep}$  de acordo com o efeito médio estimado pelo modelo e valores amostrais das variáveis para cada entrada país-ano do banco de dados.

$$Y_{it}^{rep} = \text{Bernoulli}(p_{it}),$$

em que:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{renda} * renda_{it} + b_{educ} * educ_{it} + b_{patente} * patente_{it} + b_{openk} * openk_{it} + b_{renda.educ} * renda_{it} * educ_{it})$$

A taxa de acerto média do modelo descrito acima foi de 83%. A taxa média de acerto de verdadeiro-positivos, definida pelo total de acertos nas previsões de adesão em relação ao total de adesões, é de 66%, e a taxa média de verdadeiro negativo, definida pelo total de acertos nas previsões de não-adesão em relação ao total de adesões, é de 90%. Esses números indicam que o modelo acerta bastante para prever não adesão, porém erra um pouco mais na previsão de adesão. Isso é de certa forma esperado, na medida em que temos muito menos adesão - 13% dos dados são de adesão.

Se compararmos esse resultado com um modelo ingênuo que tomasse como probabilidade de adesão 13%, ou seja, a proporção de países participantes, a taxa média de acerto seria de 23% e a taxa de verdadeiro-positivo de 13%.

Temos evidência, portanto, de que nosso modelo se ajusta bem aos dados. Para levar em conta a incerteza, mais uma vez utilizamos a *posterior predictive checks*. Denotando os dados simulados para a *posterior predictive check* por  $Y_{ij}^{rep}$  (e não mais  $Y_{.i}^{rep}$ ), temos:

$$Y_{ij}^{rep} = \text{Bernoulli}(p_{itj}),$$

em que:

$$p_{it} = \text{logit}^{-1} (\alpha_t + b_{rendaj} * \text{renda}_{it} + b_{educj} * \text{educ}_{it} + b_{patentej} * \text{patente}_{it} + b_{openkj} * \text{openk}_{it} + b_{renda.educj} * \text{renda}_{it} * \text{educ}_{it})$$

O índice  $j$  indica uma realização da densidade a posteriori para o parâmetro.

Os gráficos 6.12a e 6.12b apresentam as taxas de acerto e de verdadeiro-positivo do modelo do *posterior predictive check*. No painel esquerdo (6.12a) vemos que a taxa de acerto é maior que 75% e se concentra em torno de pouco mais de 80%. No painel direito (6.12b), a taxa de verdadeiro positivo é menor e com variabilidade maior, porém consistentemente maior do que a taxa de acerto do modelo ingênuo, reportada anteriormente.

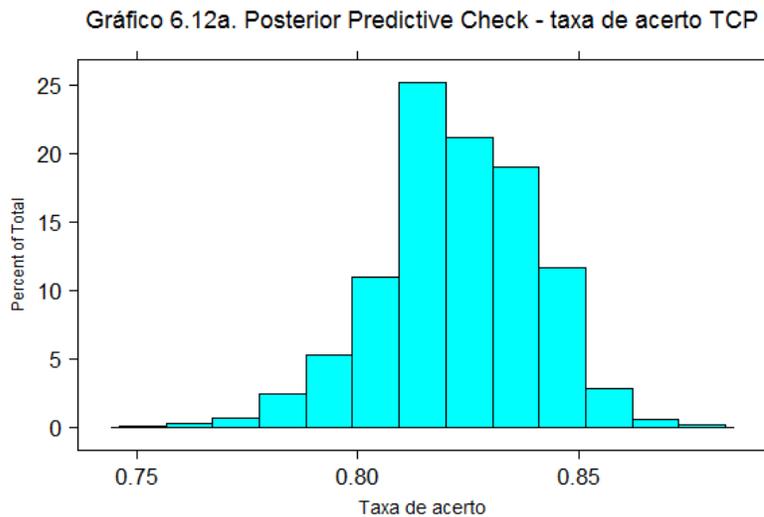


Gráfico 5.12a. O gráfico apresenta o histograma da taxa de acerto do modelo para cada um dos valores amostrados por meio de MCMC para os parâmetros do modelo. Pelo histograma, percebe-se que a taxa de acerto está entre 77% e 87%.

Gráfico 6.12b Posterior Predictive Check - verdadeiro positivo - TCP

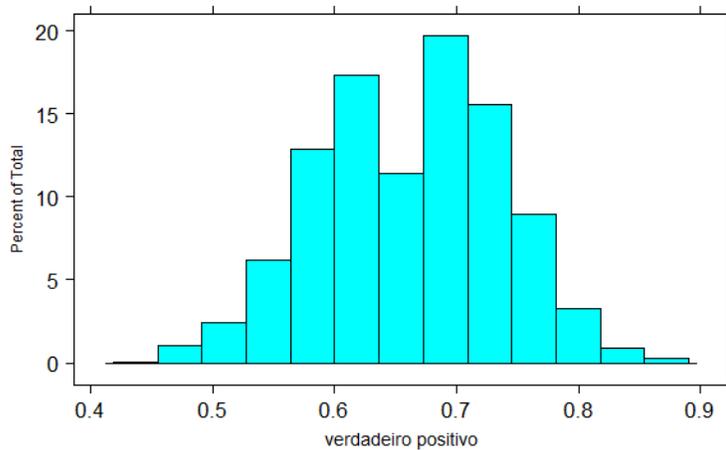


Gráfico 5.12b. O gráfico apresenta o histograma da taxa de verdadeiro positivo do modelo – prevê que o estado aderiu e de fato o estado aderiu - para cada um dos valores amostrados por meio de MCMC para os parâmetros do modelo. Pelo histograma, percebe-se que a taxa de acerto está entre 50% e 85%.

Não obstante o bom ajuste do modelo, ele comente alguns erros, especialmente no que tange à previsão de sucessos e é importante saber onde o modelo está errando. O gráfico 6.12c permite visualizar para quais países da amostra o modelo erra e quais países acerta, além do período de tempo.

## 7. Discussão dos Resultados e Considerações Finais

No presente trabalho estudamos os determinantes da adesão à Convenção de Paris e ao Tratado de Cooperação de Patentes por meio de um modelo hierárquico Bayesiano. Os resultados dos dois tratados diferem um pouco.

De um lado, os benefícios diretos da adesão ao tratado, medidos pelo número de patentes nos EUA, são significativos e importantes determinantes da adesão a ambos os tratados. Estados com mais patentes no exterior (EUA) têm maior probabilidade de adesão à Convenção de Paris. Esse resultado sugere que os países que mais obtêm patentes por meio de inovação têm mais a ganhar com a Convenção de Paris. De outro lado, países com menor nível de inovação tecnológica tendem a perder com a Convenção de Paris e, por isso, buscaram adiar a participação no tratado. As evidências apontando o efeito temporal positivo sugerem também que essa estratégia de não aderir à Convenção de Paris era mais forte antes dos anos 90. Acreditamos que não é coincidência o fato de que nos anos 90 observamos o colapso do socialismo real e a criação de um regime de patentes mais forte, com o acordo TRIPs na OMC.

Similarmente para o TCP, quanto maior o número de patentes nos EUA, maior a probabilidade de adesão. Esse é um resultado esperado, já que o objetivo explícito do TCP é reduzir o custo de transação para obtenção de patentes no exterior. E os países com mais a ganhar são aqueles com maior número de patentes externamente.

De outro lado, o custo de adesão, medido pelo índice Ginarte-Park é um preditor importante para a Convenção de Paris, mas não para o TCP. No caso do TCP, que é um tratado sobre meios, ou seja, que visa reduzir custos de transação, nossos resultados contrariam a literatura sobre *compliance*. Na medida em que é um tratado com poucos custos de adesão, estes não são significativos e são os benefícios que determinam a adesão ou não ao tratado.

Já os resultados da Convenção de Paris corroboram a literatura sobre *compliance*. Países sem legislação de patentes ou com pouca proteção de patentes terão maior custo de cumprir o tratado do que países com legislação mais rígida. O caso brasileiro, mais uma vez, é ilustrativo. Segundo Gontijo (2005), o governo brasileiro se recusou a adotar a Convenção de Estocolmo até os anos 90, entre outras razões, pela restrição que a Convenção demandava (ainda que limitada) em termos de política industrial na área de inovação. Assim, ou o Brasil

teria de mudar sua política, ou bem teria de mudar a legislação. Em ambos os casos, um movimento custoso politicamente dentro do desenvolvimentismo predominante no país.

Esse resultado, contudo, não é de todo óbvio *ex-ante*, na medida em que um país com legislação patentária fraca poderia sinalizar para a comunidade internacional desejo de mudança ou adesão aos padrões internacionais. Assim, teria mais a ganhar com a adesão do que um país que já tinha uma legislação patentária rígida. Mais uma vez, vale a pena a esse respeito retomar o caso Brasileiro. Segundo Gontijo (2005), a adesão do Brasil à Convenção de Paris em 1883 ajudou a moldar a própria legislação doméstica, já que previamente a legislação patentária era bastante fragmentada. Definições, conceitos, direitos de marcas, modelos de utilidades e patentes, todos são fatores importantes a incluir numa legislação. E países sem legislação na área ou com legislação incipientes, mas desejosos de alterar o quadro poderiam ver na Convenção de Paris uma forma de sinalizar de maneira crível essa alteração na legislação doméstica.

Renda per capita é altamente significativo para adesão ao TCP e apenas marginalmente significativo para a Convenção. Países com maior nível de renda per capita têm um mercado com maior número de patentes de estrangeiros e maior demanda das empresas por patentes no exterior. Assim, quanto maior a renda, maior o benefício de participar do tratado, levando à maior probabilidade de adesão. Apenas pelo tipo de tratado não é possível explicar a diferença de significância.

Um achado contra-intuitivo é que quanto maior o nível educacional, menor a probabilidade de adesão à Convenção de Paris. Porém, devido ao possível viés introduzido na análise, não é possível identificar os limites superiores dos parâmetros. Portanto, esse resultado carece de mais evidências empíricas.

Nosso estudo também achou evidências de que países com menor nível de escolaridade e menor nível de abertura comercial aderem mais ao TCP. Nossa interpretação desses resultados é que países com menor nível de escolaridade e abertura comercial procuram aderir ao TCP para obter efeitos colaterais como sugere a literatura para outros tratados (Hathaway, 2007). Nosso estudo apresentou evidências também de que para países mais desenvolvidos, onde os benefícios colaterais são menores, o efeito negativo da educação é menor e/ou inexistente, corroborando a teoria dos benefícios colaterais, já que países desenvolvidos teriam pouco a ganhar indiretamente da adesão ao tratado.

Além disso, temos as demais variáveis que sugerem consequências colaterais. Abertura comercial tem efeito negativo e é apenas marginalmente significativo para predizer adesão ao TCP. FDI tem efeito negativo, mas pouco significativo, para predizer adesão à Convenção.

Contrariamente à literatura sobre Direitos Humanos, regime político não é significativo para nenhum dos tratados. Em outras palavras, nossa amostra não permitiu identificar o efeito do regime político sobre adesão aos acordos estudados.

E em ambos os acordos identificamos um efeito temporal importante. Nós interpretamos esse resultado como um indicador da importância dos acontecimentos no nível sistêmico nos anos 90, particularmente o fim da Guerra Fria e a vitória ideológica do Capitalismo sobre o Socialismo real e o fortalecimento do regime de patentes com a emergência do acordo TRIPs na OMC.

Para explicar as diferenças nos achados entre a Convenção de Paris e o TCP, acreditamos que o conteúdo do tratado é um fator importante. Baseado na literatura que enfatiza o problema que um regime procura resolver (Rittberger, 1993), a Convenção de Paris foi classificada como um tratado sobre valores e o TCP como um tratado sobre meios. Outra possibilidade de classificação, ainda com base na literatura, é considerar a classificação dos tipos de jogos. Nesse caso, a Convenção de Paris seria um jogo dilemático. E o TCP seria um jogo de coordenação, em que o mais importante é a adoção de um padrão comum.

Tratados sobre valores, que implicam em redistribuição de custos e ganhos entre vencedores e perdedores domesticamente tendem a tornar os custos de adesão ao tratado uma variável importante. Por ser esse o caso da Convenção de Paris, o índice Ginarte-Park foi uma variável significativa. Países com maior grau de proteção de patentes não somente têm menor custo de adesão, como também têm os valores alinhados com o objetivo do acordo. Estados com menor grau de proteção têm de abdicar de estimular a imitação sem pagamento de royalties como forma de adoção de tecnologia. Assim, têm maior custo de adesão e valores menos alinhados.

No caso do TCP, o fator custo de adesão não é relevante, na medida em que se trata de um tratado que visa a reduzir custos de transação. Nesse caso, ganhos diretos e benefícios colaterais são variáveis mais importantes para prever a adesão.

Países pouco desenvolvidos, com baixo nível educacional e baixo nível de renda, embora com pouco a ganhar diretamente, podem aderir ao TCP para obter outros benefícios, já que o custo de adesão é quase nulo.

Interpretados dessa forma, nossos resultados sugerem a importância de incorporar interação entre variáveis sistêmicas e domésticas para entender adesão a acordos internacionais. Estrutura temporal, tipo de tratado e fatores domésticos interagem para prever adesão ao regime de patentes nos casos estudados.

Tradicionalmente, a literatura de regimes se concentrou em variáveis sistêmicas para explicar adesão a regimes internacionais. A literatura mais recente de regimes internacionais, preocupada especificamente com a adesão a tratados internacionais, tem procurado investigar a importância de fatores domésticos e sistêmicos (Bernauer, Kalbhenn, Koubi e Spilker, 2010; Baccini e Koenig-Archibugi, 2010; Hathaway, 2007; Frank, 1999). Nosso trabalho avança essa agenda ao estudar um regime ainda não analisado pela literatura: o regime de patentes.

Assim, a principal contribuição desse trabalho é avançar a agenda de pesquisa de regimes internacionais ao enfatizar os aspectos domésticos e sua interação com variáveis sistêmicas na explicação da adesão a tratados internacionais. Além disso, preenchemos uma lacuna da literatura, que não havia estudado ainda o regime de patentes.

Outros tratados de patentes existem e mesmo do regime mais amplo de propriedade intelectual. Assim, é necessário complementar os achados desse trabalho comparando com estudos de outros tratados do regime de propriedade intelectual, e posteriormente, com outros regimes (Direitos humanos, comércio, meio-ambiente entre outros). Para tanto, será necessário desenvolver uma tipologia de regimes, de forma a permitir a comparação dos resultados.

Nossa classificação dos tratados, baseada na literatura, não é muito desenvolvida por termos apenas dois casos. O teste de hipóteses requer mais observações para ser feito. Porém, os estudos de casos podem sugerir hipóteses e caminhos de investigação futuros.

## 8. Referências Bibliográficas

ABRAMOVITZ, M. (1956). "Resource and Output Trends in the United States Since 1870." *American Economic Review*, Papers and Proceedings of the American Economic Associations, 46, pp. 5-23.

ACEMOGLU, Daron; AGHION, Philippe; ZILIBOTTI, Fabrizio. (2002). "Distance to Frontier, Selection and Economic Growth." Acessado em agosto de 2009. Disponível em: <http://www.nber.org/papers/w9066>.

ACEMOGLU, Daron; JOHNSON, Simon; ROBINSON, James A.; YARED, Pierre. (2005). "From Education to Democracy?" *American Economic Review*, 95(2): 44-49.

AXELROD, Ronert; KEOHANE, Robert O. (1986). "Achieving Cooperation under Anarchy: Strategies and Institutions". In Kenneth Oie (ed.), *Coperaton Under Anarchy*, Princeton University Press.

BACCINI, Leonardo; KOENIG-ARCHIBUGI, Mathias. (2010). "Why Do States Commit to International Labour Standards? The Importance of 'Rivalry' and 'Friendship'." *Working Paper*. The London School of Economics and Political Science, Political Science and Political Economy Group.

BANCO MUNDIAL. (2010). World Bank indicators. Disponível em: <HTTP://data.worldbank.org/indicator>. Acessado em abril de 2011.

BARRO, Robert J.; LEE, Jong-Wha. (2000). "International Data on Educational Attainment: Updates and Implications." Harvard University. *CID Working Paper* No. 42, april.

BECK, N.; KATZ, J. N.; TUCKER, R. (1998). "Taking Time Seriously: Time-Series Cross-Section Analysis with a Binary Dependent Variable." *The American Journal of Political Science Review*, vol. 21, n. 4, pp 489-493.

BERNARDO, J. M. (1994). "Bayesian estimation of political transition matrices." *Statistical Decision Theory and Related Topics 5* (S. S. Gupta and J. O. Berger, eds). New York: Springer, 135-140.

BERNAUER, Thomas; KALBHEN, Anna; KOUBY, Vally; SPILKER, Gabirel. (2010). "A Comparison of International and Domestic Sources of Global Governance Dynamics." *British Journal of Political Science*, vol. 40, pp. 509-538.

BOLDRIN, Michele; LEVINE, David K. (2007). "Against Intellectual Monopoly", Cambridge University Press.

BRAGA, P.; FINK, C. (1998). "The Relationship between *Intellectual Property Rights* and Foreign Direct Investment, *Duke*." *Journal of Comparative and International Law*, vol. 19, pp. 163-187.

BRAMBOR, T; CLARK, W. R. & GOLDER, M. (2006). "Understanding Interaction models: improving empirical analysis." *Political Analysis*, 14, pp. 63-82.

- Brooks, S. P. and Gelman, Andrew. (1997). "General methods for monitoring convergence of iterative simulations". *Journal of Computational and Graphical Statistics*, vol. 7, 434-455.
- CHANG, H. (2004). "Chutando a Escada: a estratégia do desenvolvimento em perspectiva histórica." ed. Unesp.
- CHEN, Yongmin; PUTTITANUM, Thitima. (2005). "Intellectual property rights and innovation in developing countries." *Journal of Development Economics*, 78, pp. 474-493.
- CHEIBUB, José Antonio; GANDHI, Jennifer; VREELAND, James Raymond. (2009). "Democracy and dictatorship revisited." *Springer Science+Business Media*. Disponível em: [https://netfiles.uiuc.edu/cheibub/www/DD\\_page\\_files/Cheibub%20Gandhi%20Vreeland%20DD%20Revisited.pdf](https://netfiles.uiuc.edu/cheibub/www/DD_page_files/Cheibub%20Gandhi%20Vreeland%20DD%20Revisited.pdf). Acessado em 24 de agosto de 2010.
- COOPER, A. F.; HIGGOTT, R. R. A.; NOSSAL, K. R. (1993). "Relocating Middle Powers", University of British Columbia Press, Vancouver.
- DAI XY. (2002). "Political regimes and international trade: The democratic difference revisited." *American Political Science Review*, 96(1): 159-65.
- DIWAN, Ishac; RODRIK, Dani. (1991). "Patents, appropriate technology, and North-South trade." *Journal of International Economics*, Volume 30, Issues 1-2, February 1991, Pages 27-47
- DRAHOS, Peter. (1998). "Universality of Intellectual Property Rights". Disponível em: <http://www.wipo.int/tk/en/hr/paneldiscussion/papers/pdf/drahos.pdf>. Acessado em janeiro de 2010.
- FALVEY, Rod; FOSTER, Neil; GREENAWAY, David. (2006). "Intellectual Property Rights and Economic Growth." *Review of Development Economics*. Volume 10, Issue 4, pages 700-719, November.
- FARBER, Daniel. (2002). "Rights and Signals." *Journal of Legal Studies* 31:83-98.
- FIORI, J. L. (2005). "Sobre o poder global." *Novos Estudos Cebrap*, n. 73, nov., pp. 61-72.
- FLOM, P. L.; STRAUSS, S. M. (2003). "Some Graphical Methods for Interpreting Interaction Terms in Logistic and OLS Regressions." *Multiple Linear Regression Viewpoint*. 29(1), pp. 1-7.
- FRANK, David John. (1999). "The Social Bases of Environmental Treaty Ratification, 1900-1990". *Sociological Inquiry*, vol.69, n. 4, pp. 523-550.
- GAUBATZ, Kurt Taylor. (1996). "Democratic States and Commitment in International Relations."
- GAUSS, C.F. (1809). "Theoria Motus Corporum Coelestium." Hamburg: Perthes et Bresser.
- GAUSS, C.F. (1823). "Theoria Combinationis Observationum Erroribus Minimis Obnoxiae." Göttingen: Königlichen Gesellschaft der Wissenschaften.

- GELMAN, Andrew. (2003a). “A Bayesian Formulation of Exploratory Data Analysis and Goodness-of-fit Testing”. *International Statistical Review*, vol. 71, n.2, pp. 389-382.
- GELMAN, Andrew. (2003b). “Exploratory Data Analysis for Complex Models”. *Journal of Computational and Graphic Statistics*, vol. 13, n. 4, pp. 755/779.
- GELMAN, Andrew. (2004). “Exploratory Data Analysis for Complex Models.” *Journal of Computational and Graphical Statistics*, Volume 13, Number 4, Pages 755–779. Disponível em: <http://www.stat.columbia.edu/~gelman/research/published/p755.pdf>. Acessado em julho de 2011.
- GELMAN, Andrew. (2008). “Objections to Bayesian Statistics”. *Bayesian Analysis*, vol. 3, n. 3, pp. 445-450.
- GELMAN, A. (2011). “Useful models, model checking, and external validation: a mini-discussion.” Disponível em: [http://www.stat.columbia.edu/~cook/movabletype/archives/2010/08/useful\\_models\\_m.html](http://www.stat.columbia.edu/~cook/movabletype/archives/2010/08/useful_models_m.html). Acessado em julho de 2011.
- GELMAN, Andrew; CARLIN, John B.; STERN, Hal S. e RUBIN, Donald B. (2004). *Bayesian Data Analysis*. Chapman e Hall, Wahsington, segunda ed.
- GELMAN, Andrew; HILL, Jennifer. (2007). “Data Analysis Using Regression and Multilevel/Hierarchical Models.” New York: Cambridge University Press.
- GELMAN, Andrew; MECHELEN, Iven Van; VERBEKE, Geert; HEITJAN, Daniel F.; MEULDERS, Michel. (2005). “Multiple Imputation for Model Checking: Completed-Data Plots with Missing and Latent Data.” *Biometrics* 61, 74–85, March. Disponível em: [http://www.stat.columbia.edu/~gelman/research/published/biom\\_031010.pdf](http://www.stat.columbia.edu/~gelman/research/published/biom_031010.pdf). Acessado em Fevereiro 2011.
- GELMAN, Andrew; MENG, Xiao-Li; STERN, Hal. (1996). “Posterior Predictive Assessment of Model Fitness via Realized Discrepancies”. *Statistica Sinica*, 6, pp. 733-807.
- Gelman, A and Rubin, Donald. B. (1992). “Inference from iterative simulation using multiple sequences”, *Statistical Science*, vol. 7, 457-511.
- GELMAN, A.; SHALIZI, Cosma. (2010). “Philosophy and Practice of Bayesian Statistics”. Disponível em: <http://www.stat.columbia.edu/~gelman/research/unpublished/philosophy.pdf>. Acessado em Janeiro de 2011.
- GELMAN, A.; TUERLINCKX, F. (2000). “Type S error rates for classical and Bayesian single and multiple comparison procedures.” *Computational Statistics* 15:373–390.
- GELMAN, Andrew; WEAKLIEM, David. (2009). “Of Beauty, Sex and Power: too little attention has been paid to the statistical challenge in estimating small effects.” *American Scientist*, vol. 97, julho-agosto, pp. 310-316.
- GILL, Jeff. (1999). “The in significance of Null Hypothesis Significance Testing.” *Political Research Quarterly*, 52: 647-674.

- GILL, Jeff. (2008). "Bayesian Methods. A Social and Behavioral Sciences Approach." London: Chapman & Hall/CRC.
- GILL, Jeff; WALKER, Lee D. (2005). "Elicited Priors for Bayesian Model Specifications in Political Science Research". *Journal of Politics*, vol. 67, n. 3, pp. 841, 872.
- GILPIN, Robert. (1987). "The political economy of the international relations". Princeton: Princeton University Press.
- GINARTE, Juan C.; PARK, Walter G. (1997). "Determinants of Patent Rights: A Cross-National Study." *Research Policy*, October, 26(3), pp. 283-301.
- GLAESER, Edward L.; PONZETTO, Giacomo A. M.; SHLEIFER, Andrei. (2007). "Why does democracy need education?" *Journal of Economic Growth*, vol. 12, n. 2, pp. 77-99.
- GONTIJO, Cícero. (2005). "As Transformações do Sistema de Patentes, da Convenção de Paris ao Acordo TRIPS: a posição brasileira." Disponível em: <http://fdcl-berlin.de/publikationen/fdcl-veroeffentlichungen/fdcl-cicero-gontijo-as-transformacoes-do-sistema-de-patentes-maio-2005/>. Acessado em abril de 2010.
- GOULD, D.M; GRUBEN, W.C. (1996). "The role of intellectual property rights in economic growth." *Journal of Development Economics* 48 (March), 2: 323–350.
- GRIMMER, Justin. (2011). "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases". *Political Analysis*, vol. 18, n. 1, pp. 1-35.
- GROSSMAN, Gene M.; LAI, Edwin L.-C. (2004). "International Protection of Intellectual Property" (May 2004). *CESifo Working Paper Series No. 790*. Disponível em: <http://ssrn.com/abstract=348105>
- HASENCLEVER, Andreas; MAYER, Peter; RITTBERGER, Volker (1997). "Theories of International Regimes." Cambridge University Press.
- HATHAWAY, Oona A. (2007). "Why do Countries Commit to Human Rights Treaties". *Journal of Conflict Resolution*, vol. 51, n. 4, pp 588-621.
- HESTON, Alan; SUMMERS, Robert; ATEN, Bettina. (2009). "Penn World Table Version 6.3", Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, August.
- HOLLYER, James R.; ROSENDORF, Peter. (2011). "Why Do Authoritarian Regimes Sign the Convention Against Torture? Signaling, Domestic Politics, and Non-Compliance with B." *Quarterly Journal of Political Science*. forthcoming, 2011
- HU, Albert G. Z.; JAFFE, ADAM B. (2007). "IPR, innovation, economic growth and development." Department of Economics, National University of Singapore. Disponível em: [https://www0.gsb.columbia.edu/ipd/pub/Hu\\_Jaffe.pdf](https://www0.gsb.columbia.edu/ipd/pub/Hu_Jaffe.pdf). Acessado em 23 de agosto de 2010.
- HUBBARD, R.; LINDSAY, R. M. (2008). "Why P Values Are Not a Useful Measure of Evidence in Statistical Significance Testing." *Theory of Psychology*. 18 (1) , pp. 69-88.

- HUGHER, James W.; MOORE, Michael J.; SNYDER, Edward A. (2002). "Napsterizing' Pharmaceuticals: Access, Innovation, and Consumer Welfare." *Working Paper 9229* de setembro de 2002. Disponível em <http://www.nber.org/papers/w9229>. Acessado em 21/04/2005.
- JACKMAN, Simon. (2009). "Bayesian Analysis for the Social Sciences". Wiley.
- JACKMAN, Simon. (2004). "Bayesian Analysis for Political Research". *Annual Review of Political Science*, vol. 7, pp. 483-505 Wiley.
- JACKMAN, Simon. (2000). "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo". *American Journal of Political Science*, vol. 44, n. 2, pp. 375-404.
- KANWAR, Sunil. (2010). "Intellectual property protection and the licensing of technology to developing countries." *Working paper* n. 188, Center for Development Economics. Disponível em: <http://www.cdedse.org/pdf/work188.pdf>. Acessado em 23 de agosto de 2010.
- KEOHANE, Robert O. (1969). "Institutionalization in the United Nations General Assembly." *International Organization* (Autumn): 859-896.
- KEOHANE, Robert O. (1984). "After Hegemony: Cooperation and Discord in the World Political Economy". Princeton University Press.
- KEOHANE, Robert O. (1989). "Neoliberal institutionalism: A perspective on World Politics". In *International Institutions and State Power: Essays in International Relations Theory*, Boulder, Colo. Westview Press.
- KEOHANE, Robert O. (1993). "The Analysis of International Regime: Towards a European-American Research Programme". in Volker Rittberger e Peter Mayer, eds., *Regime Theory and International Relations* (Clarendon Press, Oxford, 1993), pp. 282-314.
- KEOHANE, Robert O.; NYE Jr., Joseph S. (1977). "Power und Interdependence: World Politics in Transition". Boston. MA: Little. Brown.
- KESSLER, Carol. (2006). "Post-Cold War Effectson the Non-Proliferation Regime". *Problems of Post-Communism*, vol. 53, n.2, pp. 30-38
- KINDLEBERGER, Charles. (1973). "The world in depression, 1929-39". Berkeley: University of California Press.
- KINDLEBERGER, Charles. (1981). "Dominance and Leadership in the International Economy: Exploitation, Public Goods, adn Free Riders." *International Studies Quaterly*, 25, pp. 242-54.
- KOOP, Gary; POIRIER, Dale J.; TOBIAS, Justin L. (2007). "Bayesian Econometric Methods." Cambridge University Press.
- KRASNER, Stephen D. (1982). "Structural causes and regime consequences: regimes as intervening variables". *International Organization*, vol. 36, n.2, pp. 185-205.

- KRASNER, Stephen D. (1983). "International Regimes". Ithaca. NY: Cornell University Press.
- LAPLACE, Pierre-Simon. (1774). "Mémoire sur la probabilité des causes par les évènements." Mémoires de mathématique et de physique présentés à l'Académie royale des sciences, par divers sçavans & lûs dans ses assemblées 6:621-656. Reprinted in Oeuvres complètes de Laplace. Gauthier-Villars: Paris.
- LAPLACE, Pierre-Simon. (1781). "Mémoire sur la probabilités." Mémoire de l'academie royale de sciences de Paris 1778, 227-332.
- LEE, Jeong-Yeon; MANSFIELD, Edwin. (1996). "Intellectual Property Protection and U.S. Foreign Direct Investment." *The Review of Economics and Statistics* 78 (May), 2: 181–186.
- LERNER, Josh. (2000). "150 Years of Patent Protection". *NBER Working Paper* No. W7478. Disponível em: SSRN: <http://ssrn.com/abstract=227594>
- LEVY, Marc. A.; YOUNG, Oran R.; ZÜRN, Micahel. (1995). "The Study of International Regimes". In *European Journal of International Relations*, vol. 1, no. 3, pp. 267-330.
- LIMA, Maria R. S. (2005). "A Política Externa Brasileira e os Desafios da Cooperação Sul-Sul." *Revista Brasileira de Política Internacional*, v. 48, n. no. 1, p. 24-59.
- LYONS, Russel. (2011). "The Spread of Evidence-Poor Medicine via Flawed Social-Network Analysis," *Statistics, Politics, and Policy*: Vol. 2: Iss. 1, Article 2.
- MANSFIELD, Edward D; PEVEHOUSE, Jon. (2006). "Democratization and International Organizations." *International Organization* 60, forthcoming.
- MANSFIELD, Edward D.; MILNER, Helen V. ;ROSENDORFF, B. Peter (2000). "Free to Trade: Democracies, Autocracies and International Trade." *American Political Science Review* 94(2):305-322.
- MANSFIELD, Edward D.; MILNER H. D.; ROSENDORFF B. P. (2002). "Replication, realism, and robustness: Analysing political regimes and international trade." *American Political Science Review* 2002, March; 96(1): 167-9.
- MANSFIELD, E. D.; REINHARDT, E. (2008). "International Institutions and the Volatility of International Trade." *International Organization*; 62(4): 621-52.
- MANSKI, Charles F. (2009). "Identification for Prediction and Decision." Cambridge, Massachusetts: Harvard University Press.
- MASKUS, 2000. Keith Maskus, 2000. "Regulatory Standards in the WTO," *Working Paper* Series WP00-1, Peterson Institute for International Economics.
- McCALMAM, P. (2001). "Reaping What You Sow: An Empirical Analysis of International Patent Harmonization," *Journal of International Economics*, vol. 55, 161-181, October
- MORAVCSIK, Andrew. (2000). "The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe." *International Organization* (Spring 2000). Disponível em: <http://www.princeton.edu/~amoravcs/library/origins.pdf>. Acessado em janeiro 2011.

- MORTON, Rebecca B.; WILLIAMS, Kenneth C. (2010). "Experimental Political Science and the Study of Causality: From Nature to the Lab." Cambridge University Press.
- OLSON, Mancur jr. (1965). "The Logic of Collective Action: Public Good and The Theory of Goods". Cambridge, Mass. Harvard University Press.
- PANG, Xun. (2010). "Modeling Heterogeneity and Serial Correlation in Binary Time-Series Cross-Sectional Data: A Bayesian Multilevel Model with AR(p) errors". *Political Analysis*, vol. 18, n. 4, pp. 470-498.
- PAULINO, Carlos Daniel; TURKMAN, M. Antónia Amaral; MURTEIRA, Bento. (2003). "Estatística Bayesiana". Fundação Calouste Gulbenkian, Lisboa.
- PARK, Walter G. (2008). "International Patent Protection: 1960-2005." *Research Policy*, vol. 37, pp. 761-766.
- PORTELA, Miguel; ALESSIE, Rob; TEULINGS, Coenraad N. (2006). "Measurement Error in Education and Growth Regressions," *CESifo Working Paper Series 1677*, CESifo Group Munich.
- PRZEWORSKI, A. ; ALVAREZ, M.; CHEIBUB, J. A.; LIMONGI, F. (2000). "Democracy and Development: political institutions and well-being in the world, 1950-1990". New York: Cambridge University Press, 2000.
- REINHARDT, E. (2002). "Tying Hands without a Rope: Rational Domestic Response to International Institutional Constraints," in Daniel Drezner, ed., *Locating the Proper Authorities: The Interaction of Domestic and International Institutions* (Ann Arbor: University of Michigan Press, 2002), 77-104.
- RITTBERGER, Volker. (1993). "Research on International Regimes in Germany: The Adaptative Internalization of an American Social Science Concept", in Volker Rittberger e Peter Mayer, eds., *Regime Theory and International Relations* (Clarendon Press, Oxford, 1993), pp. 3-22.
- ROMER, P. (1990). "Are Nonconvexities Important for Understanding Growth?" *The American Economic Review* 80, 97-103.
- RUGGIE, John Gerard. (1975). "International Responses to Technology: Concepts and Trends". *International Organization*. special issue. vol. 29, no. 3, Summer. pp. 557-584.
- SACKS, H. (1969). "Crisis in International Copyright: The Protocol Regarding Developing Countries". *Journal of Business Law*, p. 26.
- SHALIZI, Cosma r.; THOMAS, Andrew C. (2011). "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies". *Sociological Methods Research*, vol40, n0. 2, pp. 211-239.
- SCHERVISH, Mark J. (1995). "Theory of statistics." Springer-Verlag: New York.
- SENNES, Ricardo U. (2001). "Brasil, México e Índia na Rodada Uruguai do GATT e no Conselho de Segurança da ONU: um estudo sobre os países intermediários", Dissertação de Doutorado, datilo.

SHOR, Boris; BAFUMI, Joseph; KEELE, Luke; PARK, David. (2007). "A Bayesian Multilevel Modeling Approach to Time-Series Cross-Sectional Data". *Political Analysis*, vol. 15, n. 2, pp. 161-181.

SHRODT, Philip A. (2010). "Seven Deadly Sins of Contemporary Quantitative Political Analysis." Acessado em setembro de 2010. Disponível em:  
<http://polmeth.wustl.edu/media/Paper/Schrodt7SinsAPSA10.pdf>

SIMMONS, Beth. (2000). "International Law and State Behavior: Commitments and Compliance in International Monetary Affairs." *American Political Science Review* 94:819-835.

SKALNES, Lars S. (2006). "U.S. Statecraft in na Unipolar World". Acessado em junho de 2011. Disponível em:  
<http://www.ridgway.pitt.edu/LinkClick.aspx?fileticket=5yts22WEgsc%3D&tabid=233>.

SMITH, K. (2011). "Did IP Eat my TFP?" Disponível em: <http://networkedblogs.com/j1ABU>. Acessado em julho de 2011.

SNIDAL, Duncan. (1985). "Coordination Versus Prisoners' Dilema: Implications for International Cooperation and Regimes". *American Political Science Review*, 79, pp. 923-42.

SOLOW, R.M. (1957). "Technical change and the aggregate production function." *Review of Economics and Statistics* 39, 312-320.

STRANGE, Susan. (1987). "The persistent myth of lost hegemony". *International Organization*, vol. 41, no 4 1987.

TEIXEIRA A. A. C.; FORTUNA, N. (2010). "Human capital, r&d, trade, and long-run productivity. Testing the technological absorption hypothesis for the portuguese economy, 1960–2001." *Research Policy* 39:335–350

TOMZ, Michael. (2002). "Democratic Default: Domestic Audience and Compliance with International Agreements". Acessado em janeiro de 2011. Disponível em:  
[www.stanford.edu/~tomz/working/apsa02.pdf](http://www.stanford.edu/~tomz/working/apsa02.pdf)

USPTO (2009). USPTO United States Patent and Trade Mark Office. Disponível em:  
<http://www.uspto.gov/>. Acessado em maio de 2009.

VAN HAM, P. (1992). "**Western doctrines on East-West trade theory, history, and policy.**" St. Martin's Press in New York.

VREELAND, J. R. (2008). "Political Institutions and Human Rights. Why Dictatorships Enter into the United Nations Convention Against Torture." *International Organization*. Vol. 62, N.1.

WALT, Stephen M. (2009) "Alliances in na Unipolar World". *World Politics*, vol. 61, n.1, pp 86-120.

WALTER, Andrew. (1993). "World power and world money." Londres:Harvester Wheatsheaf, p.249.

WIPO (2009). WIPO Intellectual Property Handbook: Policy, Law and Use. Wipo publication No. 489. Disponível em: <http://www.wipo.int/about-ip/en/iprm/>. Acessado em abril de 2009.

WIPO. (2010). Patente Cooperation Treaty. Disponível em <http://www.wipo.int/pct/en/texts/articles/atoc.htm>. Acessado em junho de 2010.

WIPO. (2011). “Protecting Innovations by Utility Models”. Disponível em [http://www.wipo.int/sme/en/ip\\_business/utility\\_models/utility\\_models.htm](http://www.wipo.int/sme/en/ip_business/utility_models/utility_models.htm). Acessado em junho de 2011.

YOUNG, Oran R. (1986). International Regimes: Toward a New Theory of Institutions. *World Politics* 39 (1):104-122

YOUNG, Oran R. (1989). “International cooperation : building regimes for natural resources and the environment.” Ithaca: Cornell University Press. p. 13.

YOUNG, O. R. 1989. “The Politics of International Regime Formation - Managing Natural-Resources and the Environment.” *Ibid.* 43 (3):349-375.<Go to ISI>://A1989AF20600001

YOUNG, Oran R. (1991). “Political Leadership and Regime Formation: On the Development of Institutions in International Society.” *International Organization* 45:281-308.

ZÜRN, Michael. (1993). “Bringing the Second Image (Back) In: About the Domestic Sources of Regime Formation”. in Volker Rittberger e Peter Mayer, eds., *Regime Theory and International Relations* (Clarendon Press, Oxford, 1993), pp. 282-314.

## 9. Anexos

### Anexo A

O modelo hierárquico Bayesiano foi estimado utilizando o software *WinBugs*, versão 1.4.3, por meio do programa R. Assim, nós chamamos o *WinBugs* por meio do R, utilizando a função *bugs*, do pacote *arm*.

Para estimar o modelo, é necessário preparar as variáveis. O código abaixo mostra como preparamos as variáveis para um dos modelos estimados. É possível notar que utilizamos apenas *prioris* fracamente informativas. De fato, para parâmetros definidos sobre a reta real, utilizamos  $\theta \sim Normal(0,1000)$  e para parâmetros definidos no conjunto dos Reais não negativos, utilizamos  $\theta \sim Uniforme(0,100)$ . No *WinBugs*, trabalha-se com a precisão ao invés da variância. A precisão nada mais é do que o inverso da variância.

```
#####
```

```
### Preparando Variáveis ###
```

```
#####
```

```
# Banco de dados é dataf3
```

```
# pad é uma função que padroniza os dados
```

```
n <- dim(dataf3)[1]
n.countries <- length(unique(dataf3$Country))
n.years <- length(unique(dataf3$year))
year = as.numeric(factor(dataf3$year))
country <- as.numeric(factor(dataf3$Country))
income <- pad(dataf3$rgdpl2.mean )
democracy <- dataf3$democracy_lag
openk <- pad(dataf3$openk.mean/dataf3$pop.mean)
educ <- pad(dataf3$tot_post_sec.x_lag)
patent <- log(dataf3$patent_us.mean+1)
part.stock = c(0, unique(dataf3$part.tcp)[-7])
inc_educ = income*educ
y <- dataf3$tcp.pct_ratification
```

```
#Após preparar as variáveis, é preciso escrever o loop para o WinBugs e salvar como
#documento de texto terminação “bug”, por exemplo : “nome_do_arquivo.bug”.
```

#Abaixo mostramos um código.

```
#####
```

```
## Modelo para WinBugs, não rodar ###
```

```
#####
```

```
model {  
  for (i in 1:n){ ## loop over n of countries * n of year  
    y[i] ~ dbin (p.bound[i], 1)  
    p.bound[i] <- max(0, min(1, p[i]))  
    logit(p[i]) <- Xbeta[i]  
    Xbeta[i] <- alpha[year[i]] + b.income*income[i] +  
    b.educ*educ[i] + b.patent*patent[i] + b.openk*openk[i] +  
    b.inc_educ*inc_educ[i]  
  }  
  
  b.educ ~ dnorm (0, .0001)  
  b.income ~dnorm(0, .0001)  
  b.inc_educ ~dnorm(0, .0001)  
  b.patent ~dnorm(0, .0001)  
  b.openk ~dnorm(0, .0001)  
  
  for (j in 1:n.years){  
    alpha[j] ~ dnorm(mu.alpha[j], tau.alpha)  
    mu.alpha[j] <- g.0 + g.1*part.stock[j]  
  }  
  
  g.0 ~ dnorm (0, .0001)  
  g.1 ~ dnorm (0, .0001)  
  sigma.alpha ~ dunif(0,100)  
  tau.alpha <- pow(sigma.alpha, -2)  
}
```

```
### #####
```

```
## Voltando para o R ###
```

```
#####
```

# é preciso criar uma lista com todas as variáveis incluídas no modelo para o WinBugs

# variável é tudo aquilo que não é definido por '<->' ou '~'.

```
tcp.data_simples <- list ("n", "y", "year", "income", "educ", "patent",  
"openk", "inc_educ", "n.years", "part.stock")
```

# É recomendável definir valores iniciais para os parâmetros. Como as variáveis foram  
#escalonadas para terem valores próximos de 0 (por padronização ou logaritmo), os valores  
#dos parâmetros tendem a ser próximos de zero e, portanto, valores iniciais próximos de zero  
#não são tão ineficientes.

```
tcp.inits_simples <- function (){  
  list (alpha=rnorm(n.years), g.0=rnorm(1), b.educ=rnorm(1),  
b.income=rnorm(1), b.patent=rnorm(1),  
  b.inc_educ=rnorm(1), b.openk=rnorm(1), g.1=rnorm(1),  
sigma.alpha=runif(1))}
```

# É necessário também definir quais parâmetros devem ser salvos pelo WinBugs

# Normalmente nós pedimos para guardar todos os parâmetros

```
tcp.parameters_simples <- c ("alpha", "b.income", "b.educ",  
"b.patent", "b.openk", "b.inc_educ", "g.0",  
"g.1", "sigma.alpha")
```

# então, é possível rodar o modelo, por meio da função bugs.

# nós incluímos, além da lista de variáveis, valores iniciais para os parâmetros e quais  
#parâmetros a serem salvos, o arquivo com o loop (nesse caso, var\_slope3.bug), o diretório  
#onde está localizado do WinBugs, o número de iterações dado por n.iter, o número de  
#simulações a serem salvas, dado por n.sims e debug=T, para que o WinBugs não feche  
#rapidamente se houver algum erro.

```
tcp.time_simples <- bugs(tcp.data_simples, tcp.inits_simples,  
tcp.parameters_simples,  
model.file="D:/documentos/papers/tese  
doutoramento/scripts/var_slope3.bug",  
bugs.directory="c:/Program Files/Winbugs/WinBUGS14/",n.iter=50000,  
n.sims=2000,debug=T)
```

#Resultado no R.

Inference for Bugs-Stockholm model at "D:/documentos/papers/tese doutoramento/scripts/var\_slope3.bug", fit using WinBUGS,  
3 chains, each with 1e+05 iterations (first 50000 discarded), n.thin = 75  
n.sims = 2001 iterations saved

	Média	sd	2.5%	97.5%
alpha1	-3.3	(0.6)	-4.6	-2.2
alpha2	-2.0	(0.4)	-2.9	-1.2
alpha3	-2.2	(0.4)	-3.0	-1.4
alpha4	-2.2	(0.4)	-3.1	-1.4
alpha5	-2.7	(0.6)	-4.1	-1.6
alpha6	-0.7	(0.4)	-1.5	-0.1
alpha7	0.2	(0.4)	-0.6	1.1
b.income	0.3	(0.2)	-0.1	0.8
b.educ	-0.8	(0.3)	-1.4	-0.2
b.patent	0.3	(0.1)	0.1	0.5
b.fdi	-0.3	(0.3)	-0.9	0.1
b.inc_educ	0.0	0.1	-0.2	0.2
g.0	-3.9	1.0	-5.8	-1.9
g.1	5.2	2.1	0.7	9.0
sigma.alpha	0.8	0.5	0.1	2.1

For each parameter, n.eff is a crude measure of effective sample size, and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

DIC info (using the rule,  $pD = \bar{D} - \hat{D}$ )

$pD = 11.3$  and  $DIC = 305.8$

DIC is an estimate of expected predictive error (lower deviance is better).

## Anexo B

A diferença preditiva média ou esperada pode ser mais bem entendida por meio de um exemplo. Assim, tomando como exemplo um modelo com dois inputs (educação e renda) e quatro preditores (intercepto, educação, renda, e educação\*renda), a diferença preditiva média de um acréscimo de uma unidade na renda é dada por:

$$\delta(educ) = \text{logit}^{-1}(a + b.educ * educ + b.renda * 1 + b.renda.educ * educ * 1) \\ - \text{logit}^{-1}(a + b.educ * educ + b.renda * 0 + b.renda.educ * educ * 1)$$

$$\text{diferença preditiva média} = \frac{1}{n} \sum_1^n \delta(educ_i)$$

Esta medida permite-nos levar em consideração o efeito de uma mudança de 0 para 1 na renda para cada país da amostra, com seus níveis de educação e tomar a média dessa diferença preditiva.

O modelo com interação educação-renda estimado foi:

$$\text{Pr}(y_{i,t} = 1) = \text{logit}^{-1}(\alpha_t + b_{renda}renda_i + b_{educ}educ_i + b_{patente}patente_i \\ + b_{openk}openk_i + b_{renda\_educ}renda\_educ_i)$$

Para simplificar a notação, não escreveremos aqui o modelo para os interceptos, já que são os mesmos dos demais modelos e foram descritos anteriormente.

A diferença preditiva de um desvio padrão na renda per capita em relação à média para um dado quinquênio t, é:

$$\delta(educ, patente, openk) \\ = \text{logit}^{-1}(\alpha_t + b_{renda} * 1 + b_{educ} * educ_i + b_{patente} * patente_i + b_{openk} \\ * openk_i + b_{renda-educ} * 1 * educ_i) \\ - \text{logit}^{-1}(\alpha_t + b_{renda} * 0 + b_{educ} * educ_i + b_{patente} * patente_i + b_{openk} \\ * openk_i + b_{renda-educ} * 0 * educ_i)$$

E a diferença preditiva média da renda é dada por:

$$\text{diferença preditiva média da renda} = \frac{1}{n} \sum_1^n \delta(educ_i, patente_i, openk_i)$$

Similarmente, podemos definir a diferença preditiva média da escolaridade:

$\delta(\text{educ}, \text{patente}, \text{openk})$

$$\begin{aligned} &= \text{logit}^{-1}(\alpha_t + b_{\text{renda}} * \text{renda}_i + b_{\text{educ}} * 1 + b_{\text{patente}} * \text{patente}_i + b_{\text{openk}} \\ &* \text{openk}_i + b_{\text{renda-educ}} * 1 * \text{renda}_i) \\ &- \text{logit}^{-1}(\alpha_t + b_{\text{renda}} * \text{renda}_i + b_{\text{educ}} * 0 + b_{\text{patente}} * \text{patente}_i + b_{\text{openk}} \\ &* \text{openk}_i + b_{\text{renda-educ}} * 0 * \text{renda}_i) \end{aligned}$$

E a diferença preditiva média da educação é dada por:

$$\text{diferença preditiva média da educação} = \frac{1}{n} \sum_1^n \delta(\text{renda}_i, \text{patente}_i, \text{openk}_i)$$