

TEXTO PARA DISCUSSÃO N° 197

**ATRASSO TECNOLÓGICO, ATRASSO SOCIAL:
UMA INVESTIGAÇÃO SOBRE AS RELAÇÕES ENTRE
PRODUÇÃO CIENTÍFICO-TECNOLÓGICA E
DESENVOLVIMENTO HUMANO NO BRASIL**

**Ana Flávia Machado
Mônica Viegas Andrade
Eduardo da Motta e Albuquerque**

Maio de 2003

Ficha catalográfica

330.34(81) Machado, Ana Flávia.

M149a

2003

Atraso tecnológico, atraso social: uma investigação sobre as relações entre produção científico-tecnológica e desenvolvimento humano no Brasil. / por Ana Flávia Machado; Mônica Viegas Andrade; Eduardo da Motta e Albuquerque. Belo Horizonte: UFMG/Cedeplar, 2003.

34p. (Texto para discussão ; 197)

1. Desenvolvimento econômico – Efeito das inovações tecnológicas - Brasil. 2. Qualidade de vida – Efeito das inovações tecnológicas - Brasil. 3. Pobreza - Brasil. 4. Bem estar social - Brasil. I. Andrade, Mônica Viegas. II. Albuquerque, Eduardo da Motta e. III. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. IV. Título. V. Série.

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL**

**ATRASSO TECNOLÓGICO, ATRASO SOCIAL: UMA INVESTIGAÇÃO SOBRE AS
RELAÇÕES ENTRE PRODUÇÃO CIENTÍFICO –TECNOLÓGICA E
DESENVOLVIMENTO HUMANO NO BRASIL**

Ana Flávia Machado
Cedeplar-UFMG

Mônica Viegas Andrade
Cedeplar-UFMG

Eduardo da Motta e Albuquerque
Cedeplar-UFMG

**CEDEPLAR/FACE/UFMG
BELO HORIZONTE
2003**

SUMÁRIO

INTRODUÇÃO	6
1. TECNOLOGIA, DESENVOLVIMENTO HUMANO E A POSIÇÃO INTERNACIONAL DO BRASIL	6
2. DESCRIÇÃO DAS VARIÁVEIS	11
3. METODOLOGIA	14
3.1. O modelo Hurdle Binomial Negativo.....	14
3.1.2 Interpretação dos Coeficientes	16
3.2. Tratamento das variáveis.....	17
4. RESULTADOS	19
4.1. Análise descritiva	19
4.2. Resultados do modelo	22
4.2.1. Modelo com IDH-M.....	22
4.2.2. Modelo com indicador de desenvolvimento urbano (f1) e indicador de presença de pobreza (f2)	27
5. CONCLUSÕES PRELIMINARES.....	30
REFERÊNCIAS BIBLIOGRÁFICAS	33
ANEXO - APÊNDICE DE TABELAS.....	34

ABSTRACT

This article evaluates the relationship between scientific and technological production and the degree of human development in Brazil, on a municipal level. The variables used can be divided into seven groups: human development index, urban infrastructure, cultural infrastructure, incentives policies, education, health and scientific technological production. We used Negative Binomial Hurdle model. The results show a strong relationship between cultural infra-structure and scientific technological production, which shows that such production is associated both with the indices of human development and with those of existence of poverty.

RESUMO

Este artigo avalia a relação entre produção científica e tecnológica e grau de desenvolvimento humano no Brasil, tomando o município como unidade de observação. As variáveis utilizadas nesse estudo podem ser definidas em sete grupos: indicador de desenvolvimento humano (IDH-M), infraestrutura urbana, infraestrutura cultural, políticas de incentivos, educação, saúde e produção científico-tecnológica. Para realizar a análise econométrica, é utilizado o modelo *Hurdle* Binomial Negativo para registro de patentes e para publicação de artigos. Os resultados mostram que há uma forte relação entre a infraestrutura cultural e a produção científica e tecnológica e indicam que a produção científica e tecnológica está associada tanto a indicadores de desenvolvimento humano como a indicadores de presença e incidência de pobreza.

JEL:I3, O3, O1

INTRODUÇÃO

O objetivo deste artigo é avaliar as relações entre produção científica e tecnológica e bem-estar social no Brasil. Essa avaliação é realizada tomando como nível de análise o município a qual permite considerar as desigualdades inter e intra-regionais presentes no país.

A motivação principal deste artigo advém da publicação do *Human Development Report* 2001. Esse Relatório, intitulado *Making new technologies work for human development* (UNDP, 2001), apresenta um conjunto importante de questões, relacionando as possíveis contribuições do avanço tecnológico para o desenvolvimento humano. A discussão apresentada no Relatório instiga uma investigação do caso brasileiro.

O *Human Development Report* 2001 sugere uma associação entre capacidade tecnológica e desenvolvimento humano dos países. Essa associação é sustentada por um conjunto de interações entre o desenvolvimento tecnológico e o desenvolvimento humano. Através de múltiplos canais de influência, que se reforçam mutuamente, origina-se um circuito virtuoso entre o avanço tecnológico e o desenvolvimento humano.

Dessa associação provém a questão central a ser investigada neste artigo. Ou seja, quais as relações entre a dimensão tecnológica e a dimensão do desenvolvimento humano no país? Até onde as situações de atraso tecnológico e de atraso social estão articuladas? Os dados trabalhados nesse artigo, que desagrega no nível municipal estatísticas de produção científica e tecnológica, estatísticas de desenvolvimento humano e estatísticas de bem-estar social, podem contribuir para uma melhor compreensão do caso brasileiro e de sua localização no cenário internacional, conforme diagnóstico do *Human Development Report* 2001.

Este artigo está dividido em cinco seções. A primeira resenha as conclusões mais importantes do *Human Development Report* 2001, sistematizando as múltiplas ligações entre tecnologia e bem-estar social e identificando a posição do Brasil no cenário internacional. A segunda apresenta os dados - e suas fontes - que são utilizados no texto, indicando problemas e limitações que eles portam. A terceira seção introduz a metodologia utilizada no artigo. A quarta seção apresenta e discute os resultados encontrados. Finalmente, a quinta seção conclui o artigo.

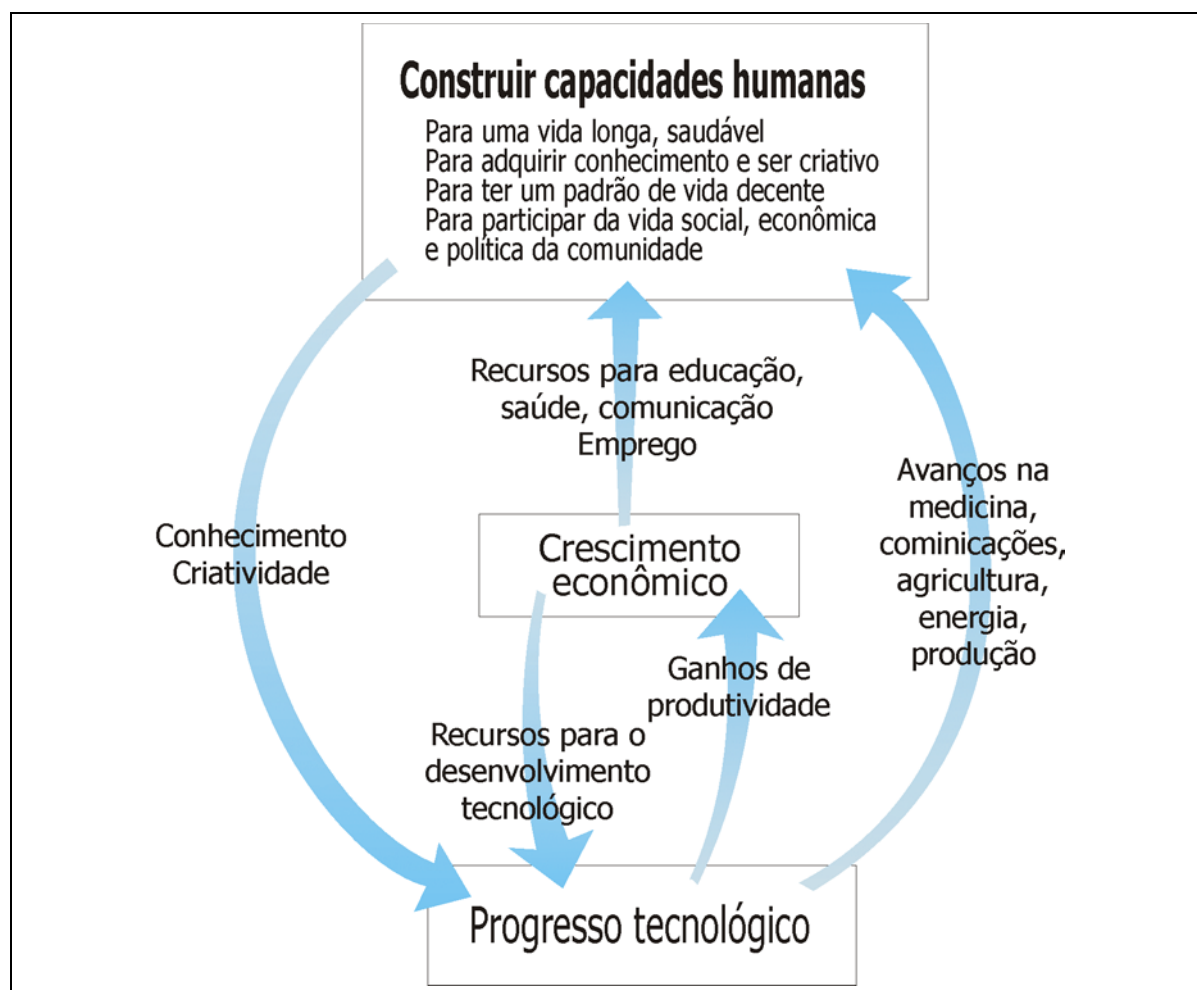
1. TECNOLOGIA, DESENVOLVIMENTO HUMANO E A POSIÇÃO INTERNACIONAL DO BRASIL

Abramovitz (1986), em uma discussão sobre os processos de *catching up*, enfatiza a importância do conceito de "capacitação social". O conceito de capacitação social aponta "...características societárias tenazes são responsáveis por uma porção, talvez uma substancial porção, do fracasso passado de um país em alcançar um nível de produtividade econômica tão alto quanto o dos países mais avançados" (p. 387). Tal conceito permite identificar que "o potencial de crescimento rápido de um país é forte não quando ele é atrasado sem qualificação, mas quando ele é tecnologicamente atrasado mas socialmente avançado" (p. 388). "Capacitação social" envolve, portanto, questões como educação, instituições financeiras, instituições políticas: "o estado da

educação incorporada na população de uma nação e os seus arranjos institucionais existentes constroem-na em sua escolha de tecnologia" (p. 388). Outros elementos da "capacitação social" seriam a ciência, a relação ciência-indústria-tecnologia e a educação geral e técnica (p. 403).

O Relatório de 2001 (UNDP, 2001) sintetiza diversos estudos, destacando a multiplicidade de canais que determinam as interações entre a tecnologia e o desenvolvimento humano. A Figura I apresenta essas relações de forma esquemática.

FIGURA I
Ligações entre tecnologia e desenvolvimento humano



Fonte: UNDP (2001, p. 28)

Esse esquema indica um conjunto de interações entre os diversos componentes envolvidos, apontando a existência de relação de causalidade em diversas direções. A Figura I mostra uma relação causal do desenvolvimento humano para o desenvolvimento tecnológico, ao mesmo tempo em que existe uma relação causal inversa. Essa relação nos dois sentidos sugere a existência de um "circuito virtuoso", de retroalimentação recíproca entre o progresso tecnológico e o desenvolvimento humano, por intermédio do crescimento econômico.

Como essas influências múltiplas se processam?

Inovações tecnológicas afetam o desenvolvimento humano de duas formas. Diretamente, disponibilizando produtos que afetam *diretamente* a saúde, a nutrição, as condições de vida das populações. Exemplos desses produtos são variedades de plantas resistentes a secas, vacinas para doenças infecciosas, energia limpa, acesso às informações disponibilizadas pela Internet etc (UNDP, 2001, p. 28). Indiretamente, por meio de inovações tecnológicas que afetam o desenvolvimento humano, em função de seu impacto sobre o crescimento econômico através de ganhos de produtividade.

Um canal importante da influência direta do avanço tecnológico sobre o desenvolvimento humano se dá através dos impactos do desenvolvimento científico-tecnológico sobre o setor saúde. Como exemplo, a UNDP afirma que “avanços médicos como imunizações e antibióticos resultaram no século XX em ganhos mais rápidos na América Latina e na Ásia Oriental do que os alcançados na Europa durante o século XIX através de melhorias na nutrição e no saneamento... Na década de 1970, a expectativa de vida nas duas regiões ultrapassou 60 anos, conseguindo em quatro décadas o que, na Europa, tendo por começo o ano de 1800, demorou um século e meio” (UNDP, 2001, pp. 28-29).

O *Human Development Report* 2001 apresenta, dentre as justificativas dessa relação entre progresso tecnológico e saúde, um estudo de Wang et al (1999). Esse estudo avalia as contribuições relativas da renda, da educação e do progresso técnico para avanços na saúde. Wang et al (1999, pp. 18-19) encontram que o progresso técnico é responsável por 45% da redução na taxa de mortalidade de menores de cinco anos de idade, por 49% da ampliação da expectativa de vida das mulheres e 50% da ampliação da expectativa de vida masculina.

A influência do desenvolvimento humano no desenvolvimento tecnológico, por outro lado, ocorre através do alcance de níveis mais elevados de educação, os quais constituem importantes fatores para a criação e difusão de inovações. O Relatório destaca, ainda, que, com o desenvolvimento humano, existe uma maior disponibilidade tanto de cientistas para assumir atividades de pesquisa como de trabalhadores em condições de aprender e dominar novas tecnologias.

Articulando esses dois sentidos causais, o Relatório sugere que “desenvolvimento humano e avanço tecnológico podem se reforçar mutuamente, criando um circuito virtuoso” (p. 28). Em relação ao progresso tecnológico, a Figura I sugere que o crescimento econômico tem relações causais nos dois sentidos. Por um lado, o crescimento econômico contribui para o progresso tecnológico, disponibilizando recursos para atividades de pesquisa, para a construção de uma infra-estrutura científico-tecnológica etc. Por outro lado, as inovações tecnológicas contribuem para ampliar a produtividade econômica.

No que tange ao desenvolvimento humano, o *Human Development Report* 2001 sugere apenas um sentido de causalidade, o efeito do crescimento econômico sobre o desenvolvimento humano. O crescimento econômico influencia o desenvolvimento humano através dos recursos disponibilizados para a educação, para a saúde, pela geração de empregos etc. Porém, outros estudos podem ser utilizados para sugerir um acréscimo na Figura I: uma seta entre o desenvolvimento humano e o crescimento econômico, indicando a existência de causalidade entre o desenvolvimento humano (ou de aspectos constitutivos desse desenvolvimento) e o crescimento econômico.

Fogel (1994), por exemplo, demonstra a influência de ganhos na nutrição sobre o crescimento da renda per capita. O Relatório do Banco Mundial de 1993, sistematiza alguns aspectos onde avanços em saúde influem diretamente sobre o crescimento econômico. Destacam-se os seguintes pontos: a) ganhos na produtividade do trabalho; b) melhor utilização de recursos naturais; c) benefícios que a educação pode trazer para próximas gerações; d) redução nos custos da assistência médica; e) influência de investimentos em saúde sobre a redução da pobreza. Somando os efeitos, é apresentada a conclusão de que “melhoria nas condições de saúde deve conduzir à melhoria do desempenho econômico no âmbito nacional” (p. 23) e que “os dados indicam que melhores condições de saúde representam crescimento mais acelerado” (p. 25).

Para a OMS (WHO, 1999), os avanços na saúde influenciam, direta e indiretamente, o crescimento econômico. Diretamente, determinando ganhos de produtividade do trabalho, e, indiretamente, através da melhoria das condições de aprendizado e de todos os efeitos decorrentes de melhor desempenho educacional.

É necessário considerar, entretanto, que o cenário internacional é caracterizado por profundas desigualdades em termos de renda (UNDP, 2001, pp. 16-20), condições de vida e desenvolvimento humano (UNDP, 2001, pp. 141-144), recursos científicos e tecnológicos (UNDP, 2001, pp. 48-51). Quanto à saúde, o cenário mundial descrito pela OMS pode ser sintetizado por um desafio duplo: “epidemias emergentes e problemas persistentes” (WHO, 1999, pp. 13-27). Esse cenário tem uma distribuição desigual, em especial na distribuição da “carga da doença evitável” (pp. 22-27). Esse conjunto de informações é importante para a discussão sobre as razões da ausência dos circuitos virtuosos entre avanço tecnológico e desenvolvimento humano em várias regiões do planeta.

No caso da saúde (um componente decisivo do desenvolvimento humano), a avaliação da situação tecnológica indica de forma dramática a enorme desigualdade existente. Diversos estudos recentes identificam de forma complementar essa desigualdade (GFHR, 2002; WHO, 1996, 2001; UNDP, 2001). Talvez a iniciativa do *Global Forum for Health Research* (GFHR, 2002) sintetize um conjunto de esforços internacionais. O *Global Forum* destaca que menos de 10% dos gastos mundiais em pesquisa na área da saúde é dedicado a doenças ou condições que representam mais de 90% da carga mundial da doença: daí o “hiato 10/90” (pp. 89-91). Uma das explicações desse hiato está na articulação entre o nível de desenvolvimento e os recursos alocados para a pesquisa e desenvolvimento em saúde: segundo a OMS, os países de baixa e média renda respondem por apenas 2,2% dos fundos globais destinados para pesquisa em saúde (WHO, 1996, p. 218).

À luz dessas discussões, buscamos identificar a posição brasileira no cenário mundial. O *Human Development Report 2001* (UNDP, 2001), contém dois indicadores interessantes de serem analisados: índice de desenvolvimento humano (IDH), e o indicador de realização tecnológica (IRT).¹ A avaliação dos dados internacionais identifica uma correlação positiva entre os países “líderes” na

¹ Esse indicador é denominado “*technology achievement index*” (TAI), calculado a partir de dados de “criação tecnológica” (patentes e receitas de *royalties*), “difusão de inovações recentes” (servidores de internet e exportações de alta tecnologia), “difusão de inovações antigas” (telefones e consumo de eletricidade”) e “habilitações humanas” (anos de escolaridade e matrículas universitárias em áreas científicas e exatas) (UNDP, 2001, pp. 46-47). Esse indicador é útil, mas contém problemas importantes. Para os objetivos gerais desse artigo ele pode ser usado, pois possibilita uma visão geral que distingue especialmente os países líderes (IRT > 0,5) e os países marginalizados (IRT < 0,2) de um grande conjunto intermediário (que poderia ser identificado como um conjunto de países com “sistemas imaturos de inovação”) onde estão incluídos os países listados na Tabela I (Brasil, Índia, México e África do Sul).

área tecnológica (países com um IRT maior do que 0,5) e os países com “alto índice de desenvolvimento humano” (IDH maior do que 0,8). Além disso, todos os países fora do conjunto dos “líderes” em tecnologia estão em grupos com IDH inferior ao considerado “alto desenvolvimento humano”.

O Brasil faz parte do conjunto de países em situação intermediária tanto em termos de desenvolvimento humano como de condição tecnológica. A Tabela I apresenta dados para Brasil, Índia, México e África do Sul. Todos esses países encontram-se em um mesmo nível de desenvolvimento humano (“médio”) e não estão entre os “líderes” em tecnologia identificados pelo Relatório.

TABELA I
Comparação entre o IDH (índice de desenvolvimento humano)
e o IRT (indicador de realização tecnológica) ²

País	Posição IDH	Valor IDH	Posição IRT	Valor IRT	Taxa de Alfabetização (% acima de 15 anos) (1999)	Índice de Gini	PNB per capita (PPP US\$) (1999)	Expectativa de vida no nascimento (anos) (1999)
Brasil	69	0,750	43	0.311	84.9	59.1	7,037	67.5
Índia	115	0,571	63	0.201	56.5	37.8	2,248	62.9
México	51	0,790	32	0.389	91.1	51.9	8,297	72.4
África do Sul	94	0,702	39	0.340	84.9	59.3	8,908	53.9

Fonte: Human Development Report (2001), elaboração própria.

É interessante notar que a ordenação entre os países em termos de IDH e ITR não se mantém. A diferença básica entre as duas classificações é a posição do Brasil e da África do Sul: enquanto a África do Sul está na frente do Brasil quanto ao IRT (respectivamente as posições 39 e 43), essa ordem é invertida quanto ao IDH (94 e 69, respectivamente).

Avaliando de forma introdutória dados sobre concentração de renda, nota-se que a África do Sul, o Brasil e o México possuem índices de Gini superiores a 50.³ A Índia parece ser um caso de distribuição mais homogênea da pobreza, pois combina o menor índice de Gini com o menor PNB per capita entre os quatro países. Do ponto de vista tecnológico, há indicações de concentração geográfica das atividades inovativas: em todos os quatro países, a unidade federativa líder detém mais de 40% do total das patentes do país (USPTO, 2002). Esse padrão de concentração geográfica repete-se nas atividades científicas, com exceção da Índia, a unidade federativa líder detém mais de 40% dos artigos científicos publicados (ISI, 2002).

Essas pistas iniciais da desigualdade interna na distribuição de renda (indicador associado ao bem estar social) e na distribuição das atividades científicas e tecnológicas estimulam a investigação

² A diferença nas posições dos países em termos dos dois indicadores (IRT e IDH) reflete a utilização de conjuntos diferentes de países. Alguns países constam da classificação do IDH mas não foram considerados no IRT: por exemplo, Suíça, Croácia, Lituânia e Latvia estão na frente do Brasil em termos de IDH e não foram avaliados em termos de IRT.

³ No Human Development Report 2001 (UNDP, 2002, p. 185), o índice de Gini é definido variando entre 0 (igualdade perfeita) e 100 (desigualdade perfeita).

mais detalhada do caso brasileiro. Essa investigação é importante porque a desigualdade interna (tanto nas atividades científicas e tecnológicas como nos indicadores de bem-estar social) pode ser um importante fator explicativo da posição internacional do Brasil (conforme indicado pelo Relatório, segundo a Tabela I).

2. DESCRIÇÃO DAS VARIÁVEIS

Buscando estimar a relação entre produção científica e tecnológica e grau de desenvolvimento humano, realizamos uma análise controlada, tendo o município como unidade de observação. Ainda que essa unidade geográfica assuma no Brasil, na maioria dos casos, estrutura produtiva de pequena escala, recorrer a informações em unidade mais abrangente, como o estado, cria problemas de heterogeneidade assim como reduz o número de observações.

As variáveis utilizadas nesse estudo podem ser definidas em sete grupos: indicador de desenvolvimento humano, infra-estrutura urbana, infra-estrutura cultural, políticas de incentivos, educação, saúde e produção científico-tecnológica.

O Índice Municipal de Desenvolvimento Humano (IDH-M) é construído por intermédio de três dimensões. A primeira é da educação, considerando dois indicadores diferentes com pesos distintos, a taxa de adultos alfabetizados (pessoas com mais de 15 anos capazes de ler e escrever) e a taxa bruta de frequência à escola. O primeiro indicador tem peso dois e o segundo possui peso um. A longevidade constitui-se em outra dimensão do IDH-M e refere-se à esperança de vida ao nascer. Por fim, a dimensão renda é a renda média de cada residente no município ou renda municipal per capita. Apenas a dimensão de longevidade não é obtida diretamente dos dados do Censo Demográfico de 2000. Recorre-se a técnicas indiretas para cálculo da estimativa de mortalidade, utilizando as respostas do Censo sobre número de filhos nascidos vivos e o número de filhos ainda vivos na data de realização do Censo para obter as proporções de óbitos. Essas proporções são transformadas em estimativas de probabilidade de morte que, posteriormente, compõem uma tabela de vida de onde são retiradas a esperança de vida ao nascer. Essa metodologia é de autoria de um grupo de pesquisadores da Fundação João Pinheiro e do IPEA e guarda algumas diferenças com relação ao IDH do PNUD no que tange à definição de renda e das variáveis de educação. As variáveis do IDH-M são transformadas em índices que variam entre 0 (pior) e 1 (melhor), e a combinação destes índices em um indicador síntese. Assim, quanto mais próximo de 1 o valor deste indicador, maior será o nível de desenvolvimento humano do município.

O grupo infra-estrutura urbana contempla as variáveis existência de favelas ou assemelhados, taxa de domicílios em áreas urbanas, taxa de domicílios com coleta de lixo, taxa de domicílios com rede de abastecimento geral de água, taxa de domicílios com banheiro. O cálculo das taxas é a relação entre os domicílios que apresentam a condição em questão e o total de domicílios no município. A variável existência de favelas ou assemelhados é uma variável binária (assume valor 1 quando há algum registro e 0 em caso contrário).

A inclusão desse grupo de variáveis permite dimensionar o papel do desenvolvimento urbano⁴ na produção de tecnologia, esperando-se uma relação positiva entre os mesmos. Além disso, como a unidade de análise é município, este tipo de controle é ainda muito importante no Brasil, em virtude da heterogeneidade no desenvolvimento econômico. Observando a média dessas variáveis, constatamos que 71% dos municípios possuem favelas ou assemelhados, apenas 59% dos domicílios se localizam em áreas urbanas, 59% dos domicílios possuem rede de abastecimento geral de água e 55% possuem coleta de lixo (vide Tabela V).

As variáveis de infra-estrutura cultural são presença de provedor de Internet e de livrarias (variável binária), número de bibliotecas públicas⁵ e número de jornais diários. Essas quatro variáveis possibilitam dimensionar o grau de acesso à informação dentro dos municípios. Quanto maior a presença de provedores de Internet, livrarias, jornais diários e bibliotecas na unidade de análise, maior a produção de patentes e de artigos.

O grupo de variáveis de políticas de incentivos está associado à capacidade produtiva do município. Programa de geração de emprego e renda, programa de capacitação profissional, existência de incentivos às atividades econômicas constituem-se nessas variáveis e são binárias.

A fonte de dados secundária é o CDROM do IBGE “Perfil de Informações Municipais” e, em sua maior parte, as informações se referem ao ano de 1999. A Pesquisa de Informações Básicas Municipais – 1999 é uma pesquisa de registros administrativos relativos às prefeituras, permitindo construir perfil abrangente dos municípios brasileiros.

O IDH-M é obtido no site do IPEA, na publicação Novo Atlas do Desenvolvimento Humano no Brasil, elaborada por pesquisadores da Fundação João Pinheiro, órgão do governo do Estado de Minas Gerais, e por pesquisadores do IPEA.

A única variável de educação é a taxa de alfabetização, parcela da população com 10 anos e mais alfabetizadas. Nesse caso, a fonte de dados é o Censo Demográfico de 2000, com informações para o universo da população. O grau de alfabetização da população reflete o potencial de capital humano do município. No Brasil, embora haja um processo recente de universalização da educação formal, a taxa de alfabetização ainda é um indicador importante quando a unidade de análise é municipal. A taxa média de alfabetização observada é de 80,66% da população de 10 anos e mais, acompanhada por um desvio padrão de 11,46 com valor mínimo de 40,9 e máximo de 99,2. (vide Tabela V)

As variáveis de saúde são taxa de mortalidade infantil e número de leitos per capita. A taxa de mortalidade infantil é calculada por meio da relação entre o número de mortos de 0 a 1 ano e o total de nascidos vivos. Para corrigir problemas de sub-registros de informações tanto de mortes quanto de nascimentos, utilizamos a série construída por Simões (1996). Buscando evitar volatilidade das taxas devido a tamanho reduzido da população, o autor calcula a taxa de mortalidade infantil para micro

⁴ Esse grupo de variáveis está associado ao PIB municipal e como não há registro desse indicador para todos os municípios brasileiros, optamos por esse grupo de variáveis.

⁵ Para a cidade de Porto Alegre, o dado original da pesquisa não está disponível. Em função disso, recorremos ao site Bibliotecas de Porto Alegre – Guia Virtual de Porto Alegre, onde apontava a presença de 18 bibliotecas. Aos demais municípios, sem informação de número de bibliotecas, não é possível imputar nenhum valor, uma vez que não existem dados disponíveis.

regiões censitárias e imputa o mesmo valor para todos municípios da mesma⁶. Essa variável é um indicador de bem-estar que complementa o acesso à educação.

Os dados sobre produção científica e tecnológica provêm de estatísticas de patentes depositadas junto ao INPI e de artigos indexados junto ao ISI. Na literatura de economia da tecnologia existem inúmeros textos que levantam as vantagens e desvantagens desses indicadores. Certamente nem todas as inovações são patenteáveis ou patenteadas, os setores industriais diferenciam-se de acordo com suas respectivas “propensões a patentear”, em países atrasados o caráter incremental e adaptativo de inovações importantes localmente pode significar que elas não são patenteadas (ver Griliches, 1990). Artigos indexados junto ao ISI não representam o conjunto da produção científica de um país, especialmente quando o país não é de língua inglesa. As disciplinas científicas também distinguem-se por diferentes “propensões a publicar”. No caso do Brasil, é importante lembrar que existe uma ampla gama de publicações científicas nacionais que não estão indexadas junto ao ISI e que as disciplinas científicas variam fortemente em termos de internacionalização da produção científica. Por isso, os dados do ISI podem ser lidos como uma espécie de “ponta do iceberg” da produção científica nacional.⁷

Apesar dos problemas, estatísticas de patentes e de artigos científicos têm contribuições para a compreensão e o mapeamento da produção científica e tecnológica do país. A recém-publicada Pesquisa sobre Inovação Tecnológica (PINTEC) do IBGE, por exemplo, ainda não tem os dados desagregados no nível de município, o que justifica a utilização dos dados de patentes como um esforço para a construção desse mapa.

As variáveis para mensurar capacidade de geração de tecnologia se referem ao número de patentes registradas no INPI entre 1988 e 1999 e ao número de artigos indexados pelo ISI em 1999. Embora no registro da patente o endereço do titular conste do documento (ver www.inpi.gov.br), o INPI não tem transferido esses dados para os seus registros magnéticos. Apenas a informação do(s) estado(s) do(s) titular(es) é apresentada.

Essa informação é importante para este trabalho, pois implica a exclusão das patentes cujos titulares são indivíduos (pessoas físicas) da presente análise. O peso dos indivíduos na patenteação do país é elevada. Em termos de patentes depositadas por residentes no Brasil, entre 1988 e 1996, de um total de 57.640 patentes, 38.802 são de pessoas físicas e 18.838 de pessoas jurídicas.

Os dados apresentados nesta seção, portanto, restringem-se a patentes cujos titulares são pessoas jurídicas. A identificação do município é realizada através do cotejamento dos dados fornecidos pelo INPI (envolvendo a identificação do titular) com as informações da RAIS (que articula o titular com um município). Do cruzamento entre as informações do INPI com a RAIS é construído o banco que é descrito nesta seção.

Os dados do INPI apresentam informações sobre 7.040 diferentes empresas, titulares de 23.919 patentes e/ou contratos de transferência de tecnologia que, entre 1990 e 2000, registram patentes e/ou averbam contratos de transferência de tecnologia. O cruzamento dessas informações com a RAIS possibilita a identificação da localização municipal de 4201 empresas, representando 17.587

⁶ No caso do município de Cuiabá, a informação provém do banco de dados do DATASUS.

⁷ Agradecemos ao Prof. Evando Mirra (CGEE-MCT) por essa observação.

patentes. Possivelmente, como o banco das patentes refere-se a um período mais amplo (1988 a 1999) e a RAIS a um único ano (1997), algumas firmas que depositam patentes no final da década de oitenta e no início da década de noventa podem ter desaparecido (por falência, por fusão, por aquisição etc). O peso de pequenas firmas no total das patentes é importante: do total das firmas presentes nos dados do INPI, 4001 (56,83% do total) depositam apenas uma patente.

O número de municípios com pelo menos um titular local de uma patente é 512. Os dez municípios com maior produção tecnológica respondem por 53,69% das patentes identificadas. É importante destacar que apenas 4 municípios (São Paulo, Rio de Janeiro, Campinas e Joinville) detêm mais de 500 patentes.

Quanto aos artigos indexados, baseiam-se em informações coletadas no site www.webofscience.fapesp.br, fornecidos pelo *Institute for Scientific Information* (ISI). Utilizando-se como critério de seleção o *Science Citation Index* (que exclui revistas da área de humanas e de artes) os 9.668 artigos publicados com autores filiados a instituições localizadas no Brasil (em 1999) são gravados a partir de consulta à Internet. A partir desta coleta, um banco de dados é construído de forma a tornar operacional o tratamento das informações relevantes (autores e suas instituições, nome da revista).

A restrição a apenas um ano (1999) é uma limitação importante da análise. Mas, dada a dificuldade da montagem do banco e a escassez de informações com esse nível de detalhe, os dados apresentados a seguir devem contribuir para esse mapeamento preliminar dos recursos científicos disponíveis no país.

O número de municípios com pelo menos um autor participante de uma instituição local é 226. Os dez municípios com maior produção científica respondem por 69% da produção científica nacional.

3. METODOLOGIA

3.1. O modelo Hurdle Binomial Negativo

Como já explicitado, o objetivo deste artigo é avaliar as relações entre produção científica e tecnológica e bem-estar social no Brasil. Embora a relação entre essas duas variáveis venha a ser em dupla direção, como mostrado na Figura 1, estamos investigando apenas em uma direção, ou seja, em que medida bem-estar social é capaz de responder pela produção científica e tecnológica no Brasil. Reconhecemos essa limitação, mas a natureza das fontes de dados utilizadas nesta pesquisa dificultam a especificação do modelo que combine ambas as direções. Temos informações para um determinado ano, não contamos com uma série temporal que nos permita avaliar os efeitos em ambas as direções.

Assim, tendo por objetivo estimar a relação entre produção científica e tecnológica e grau de desenvolvimento humano, recorreremos ao modelo *Hurdle Binomial Negativo*. Todas as variáveis do modelo estão organizadas por municípios, seguindo a divisão territorial de 1996, o que gera um número total de 5507 observações.

A estimação de modelos de determinação do número de patentes e artigos deve considerar algumas características importantes desse tipo de dado. O número de registros de patentes e de artigos publicados constituem-se em eventos, podendo os mesmos serem considerados dados de contagem (Noronha, 2001). Em primeiro lugar, observa-se que patentes e artigos publicados assumem apenas valores inteiros, não negativos e não possuem um valor limite superior. O valor associado a essas variáveis representa o número de vezes em que o evento ocorreu. Em segundo lugar, na população, quanto maior o número de patentes e artigos, menor o número de municípios que terá registrado essa frequência, o que caracteriza a distribuição desses eventos como uma distribuição Poisson.

O modelo de Poisson é o representante mais simples dos modelos de dados de contagem. Entretanto, esse modelo não é adequado à distribuição de patentes e artigos em municípios brasileiros, uma vez que a maioria dos municípios não apresenta nenhum registro. Nesse caso, a nossa distribuição não atende à hipótese básica do modelo de Poisson de equidispersão de dados, ou seja, média condicional igual à variância condicional⁸. Na literatura, existem duas formas de solucionar o problema referente ao excesso de “zeros”. Uma delas é o modelo *zero inflated* que corresponde a atribuir uma ponderação diferenciada para valores nulos e não negativos (Wooldridge, 2001). A segunda trata a sobredispersão como resultado de uma heterogeneidade não observada. Um dos modelos que considera essa solução é o Binomial Negativo, por incorporar um componente randômico no modelo de Poisson com uma distribuição gamma. Capta-se, assim, a sobredispersão, na medida em que a variância condicional passa a apresentar uma forma funcional quadrática ou linear da média condicional.

O modelo binomial negativo pode ser especificado, assumindo que $c_i > 0$ é a heterogeneidade não observada e supondo que:

$$y_i / x_i, c_i \sim \text{Poisson } c_i m(x_i \beta) \quad (1)$$

Se c_i é independente de x_i e possui distribuição gamma com variância $c_i = \eta^2$, então a distribuição de y_i dado x_i pode ser mostrada como uma distribuição Binomial Negativa com média e a variância condicional iguais a:

$$E[y_i | x_i] = m(x_i \beta) \quad (2)$$

$$V[y_i | x_i] = m(x_i \beta) + \eta^2 [m(x_i \beta)]^2 \quad (3)$$

onde:

η representa o termo de sobredispersão dos dados e y_i representa o número de patentes e artigos em cada município e x_i o vetor de covariadas.

⁸ No caso do universo total de municípios, observamos uma média de 3,47 patentes registradas por município com um desvio padrão de 72,69 e, no caso de artigos, esses números são média de 4,34 e desvio padrão de 89. A hipótese de equidispersão diz que a média condicional é igual à variância condicional. Ainda que estes números não reflitam os valores condicionados, os valores não controlados sugerem violação da hipótese.

Quando $\eta=0$, a variância é igual à média, indicando que não há sobredispersão dos dados. Nesse caso, o modelo se reduz ao modelo de Poisson. No presente estudo, estimamos uma variação do modelo Binomial Negativo, conhecido na literatura por modelo *Hurdle*. Esse modelo permite estimar o processo de decisão de registros de patentes e de publicação de artigos como dois processos estocásticos distintos⁹. O primeiro processo refere-se à decisão de produzir conhecimento e inovações e o segundo diz respeito à decisão de quantas patentes registrar e de quantos artigos publicar. Nesse caso, é razoável supor que os agentes que determinam cada um desses processos são distintos. No caso do primeiro processo, trata-se da decisão de instalação de uma instituição de pesquisa no município, da decisão pela firma de instalar um laboratório de pesquisa e desenvolvimento ou da decisão de um empresário de investir no desenvolvimento de um novo produto ou no melhoramento de um existente. No segundo processo, a quantidade de patentes registradas ou artigos publicados depende de decisão relacionada ao volume de investimentos e/ou do número de profissionais alocados para atividades inovativas.

Por esse motivo, estimamos um modelo Binomial Negativo em duas etapas, conhecido como modelo *Hurdle* Binomial Negativo. Na primeira etapa, utilizamos um modelo *logit* para determinar se o município produziu conhecimento científico e tecnológico, ou seja, para investigar os determinantes da decisão de registrar patentes e de publicar artigos. Na segunda etapa, empregamos o modelo Binomial Negativo truncado ao zero (*Negbin* truncado ao zero) para estimar o número esperado de patentes e o número esperado de artigos, considerando a amostra de municípios com geração positiva.¹⁰ Estimamos dois modelos, um para registro de patentes e o outro para publicação de artigos. Entendemos que as características do processo de produção dos mesmos são diferentes e, por isso, a relação que guardam com o desenvolvimento econômico é distinta. O que se quer investigar é a existência de diferenças nas relações entre desenvolvimento humano e a publicação de artigos e registro de patentes. Isso pode ser verificado pela análise comparativa dos coeficientes estimados.

3.1.2 Interpretação dos Coeficientes

Os coeficientes estimados são interpretados separadamente para a primeira e segunda etapa. Tanto o modelo logístico, como o modelo binomial negativo, são modelos não lineares e desse modo a interpretação dos coeficientes não pode ser feita de forma direta como usualmente é tratado no modelo linear. Em geral, estamos interessados nos efeitos marginais que mostram o efeito da variação de um dos regressores na variável dependente. A dificuldade de interpretação dos efeitos marginais em modelos não lineares é que estes dependem dos valores que as covariadas assumem na população. No caso de um modelo de estimação logístico, a forma mais fácil de interpretar os coeficientes é calcular as razões de chance. As razões de chance são usualmente calculadas para variáveis binárias. Uma razão de chance superior a 1 indica que, em relação a categoria de referência (em geral é o zero),

⁹ Como os dados são censurados, na literatura, alguns autores sugerem a estimação através do estimador Heckit, argumentando que a ausência de informação pode ser devida a um problema de seleção amostral. A estimação através do método Heckit, entretanto, desconsidera que estes dados sejam dados de contagem. Além disso, no caso de patentes e de artigos publicados, não acreditamos que o zero seja decorrente de seleção amostral, mas que represente uma escolha genuína. Desse modo, a estimação através do modelo *Hurdle* é a mais adequada.

¹⁰ O modelo *Hurdle* é estimado por intermédio do método de máxima verossimilhança, construído com duas funções parametricamente independentes. Uma função para o modelo logit tradicional e outra para o modelo binomial negativo truncado ao zero (Cameron e Trivedi, 1988).

temos uma chance maior de apresentar determinado resultado. Uma razão de chance inferior a 1 significa que temos uma chance inferior em relação à categoria de referência¹¹. Na segunda etapa do modelo, a estimação do modelo binomial negativo truncado ao zero, a interpretação dos coeficientes estimados depende do tipo de covariada que estamos utilizando, ou seja, se é uma variável binária ou contínua. O efeito marginal pode ser representado da seguinte forma: $\frac{\partial E[y/x]}{\partial x_j} = \beta_j \exp(x_i \beta)$

Essa relação mostra que o aumento (redução) no número esperado de patentes ou artigos publicados é diretamente proporcional à variação ocorrida em um dos regressores. Se a variável é contínua, o coeficiente estimado pode ser interpretado diretamente como uma semi-elasticidade, pois:

$$\begin{aligned}\frac{\partial E[y/x]}{\partial x_j} &= \beta_j \exp(x_i \beta) \\ \frac{\partial E[y/x]}{\partial x_j} \frac{1}{\exp(x_i \beta)} &= \beta_j \\ \frac{\partial E[y/x]}{\partial x_j} \frac{1}{E[y/x]} &= \beta_j\end{aligned}$$

Nesse caso, se o j-ésimo coeficiente estimado for igual a 0,12, um aumento de uma unidade na variável explicativa j implica um aumento de 12% no número esperado de patentes registradas ou artigos publicados. Para avaliar o efeito de uma variável explicativa binária considere um regressor que assume os valores 1 e 0. Nesse caso, o efeito sobre o número esperado de patentes registradas pode ser calculado como:

$$E[y/x_2, d] = \exp(\beta_1 d + x_2 \beta_2)$$

Assim:

$$\frac{E[y/x_2, d=1]}{E[y/x_2, d=0]} = \frac{\exp(\beta_1 d + x_2 \beta_2)}{\exp(x_2 \beta_2)} = \exp(\beta_1)$$

Para calcularmos o acréscimo na variável dependente quando a variável explicativa assume valor 1 em termos percentuais, efetuamos o seguinte cálculo: $[\exp(\beta)-1] \times 100$.

3.2. Tratamento das variáveis

As variáveis taxa de domicílios em áreas urbanas, taxa de domicílios com coleta de lixo, taxa de domicílios com rede de abastecimento geral de água, taxa de domicílios com banheiro, taxa de

¹¹ No caso de variáveis contínuas, o sinal positivo do coeficiente indica uma relação direta entre variações na covariada sobre a variável dependente e quanto maior o valor absoluto do coeficiente mais intensa é essa relação.

alfabetização e taxa de mortalidade infantil são altamente correlacionadas (ver matriz de correlação no apêndice). Para contornar este problema, optamos por construir índices que representem essas variáveis por intermédio do método dos componentes principais. Tal método estima fatores que são combinações lineares das variáveis originais que nos permitem explicar a variância observada na amostra.

As Tabelas II e III resumem os resultados encontrados. Os fatores 1 e 2 explicam, conjuntamente, 85% da variância total. O fator 1 contempla todas as variáveis citadas acima com distribuição uniforme dos pesos. Esse fator pode ser interpretado como indicador de desenvolvimento urbano, uma vez que todas as variáveis mostram que quanto maior o índice maior o desenvolvimento. No caso de mortalidade infantil, como essa variável entra com sinal negativo na composição do índice, municípios com uma taxa de mortalidade infantil elevada devem apresentar, portanto, um menor índice. O fator 2, por sua vez, atribui um peso maior para a mortalidade infantil, porém com sinal positivo. Desse modo, interpretamos que esse resultado estaria captando a presença de pobreza no município. Em virtude do elevado grau de explicação dos dois fatores, optamos por manter apenas os dois fatores na análise econométrica.

TABELA II
Método dos Componentes Principais Variáveis de Qualidade de Vida Urbana

Fator	Autovalor	Diferença	Proporção	Acumulado
1	3,84886	2,56458	0,64150	0,64150
2	1,28428	0,93459	0,21400	0,85550
3	0,34969	0,09602	0,05830	0,91380
4	0,25367	0,11748	0,04230	0,95610
5	0,13619	0,00888	0,02270	0,97880
6	0,12731	0,00000	0,02120	1,00000

Fonte: IBGE (2002), elaboração própria.

TABELA III
Pesos das Variáveis
Nos fatores 1 e 2

Descrição da Variável	F1	F2
Tx mortalidade Infantil	-0,34428	0,55850
Rede geral água	0,38952	0,41101
Banheiro	0,43335	-0,23101
Coleta de lixo	0,45132	0,28748
Alfabetização	0,42061	-0,42431
Urbanização	0,40167	0,45067

Fonte: IBGE (2002), elaboração própria.

4. RESULTADOS

O universo de municípios é de 5.507. Como em alguns municípios, o(s) registro(s) de algumas das variáveis utilizadas não está bem especificado, retiramos esses municípios, reduzindo o banco de dados a 4.969 observações¹².

4.1. Análise descritiva

Dos 4969 municípios, 494 apresentam registros de patentes e 204 possuem artigos publicados no ano de 1999, sendo que 562 possuem ou patentes ou artigos. Dividindo o universo de municípios entre aqueles que possuem algum tipo de produção científica e tecnológica (562) e aqueles que não possuem (4407), destacamos a diferença nas frequências de municípios com provedor de Internet e livrarias. Enquanto 64% dos municípios com produção científica possuem provedor de Internet, somente 10% dos municípios sem produção científica e tecnológica o possuem. No tocante a livrarias, 80% dos municípios com produção científica e tecnológica registram ter pelo menos uma livraria ao passo que nos municípios sem produção, esse percentual se reduz para 32%. (Veja Tabelas IV e IV-A).

TABELA IV
Estatística Descritiva para os municípios sem produção científica e tecnológica (N=4407)

Descrição da Variável	Frequência		Percentual	
	Não tem	Tem	Não tem	Tem
Favelas	3268	1139	74,15	25,85
Incentivos p/ atração de ativ. Econômica	1909	2498	43,32	56,68
Programa de Trabalho e Renda	2397	2010	54,39	45,61
Programa de capacitação profissional	2449	1958	55,57	44,43
Provedor de Internet	3963	444	89,93	10,07
Livraria	2985	1422	67,73	32,27

Fonte: IBGE (2002), INPI (2001), ISI (1999), elaboração própria.

TABELA IV-A
Estatística Descritiva para os municípios com produção científica e tecnológica (N=562)

Descrição da Variável	Frequência		Percentual	
	Não tem	Tem	Não tem	Tem
Favelas	258	304	45,91	54,09
Incentivos p/ atração de ativ. Econômica	69	493	12,28	87,72
Programa de Trabalho e Renda	164	398	29,18	70,82
Programa de capacitação profissional	160	402	28,47	71,53
Provedor de Internet	199	363	35,41	64,59
Livraria	111	451	19,75	80,25

Fonte: IBGE (2002), INPI (2001), ISI (1999), elaboração própria.

¹² Das 538 observações excluídas, nenhuma delas apresenta produção científica e tecnológica. Observando a distribuição das variáveis independentes nessa sub-amostra, constatamos que tratam-se de municípios com menor grau de desenvolvimento. Nesse sentido, a ausência de registro parece estar correlacionada com o grau de desenvolvimento, configurando viés nas observações eliminadas.

No que tange às variáveis número de bibliotecas públicas e número de jornais diários, constatamos que o conjunto dos municípios que não têm nem patentes e nem artigos (4407) possuem menos de uma biblioteca pública por município (0,92) e 2,43 jornais diários. Em contraste, os municípios com algum tipo de produção científica e tecnológica, apresentam duas bibliotecas por municípios e 6,13 jornais diários. O IDH-M médio dos municípios sem produção tecnológica e científica (0,6926) é inferior ao IDH-M médio dos municípios com esse tipo de produção, 07937 (Vejas Tabelas IV-B e IV-C).

TABELA IV-B

**Estatística das Variáveis contínuas para municípios sem produção científica e tecnológica
(Universo total / N=4407)**

Descrição da Variável	Média	Desvio Padrão	Mínimo	Máximo
Biblioteca	0,9253	0,7756	0,0000	16,0000
Jornal (diário)	2,4350	9,5983	0,0000	500,0000
Tx rede de esgoto	0,5594	0,2290	0,0000	1,0000
Banheiro	0,8186	0,1970	0,0637	1,0000
Coleta de lixo	0,5107	0,2450	0,0000	0,9984
Área urbana	0,5609	0,2186	0,0000	1,0000
Alfabetizado	79,2386	11,2709	40,9000	99,2000
IDH-M	0,6926	0,0782	0,4668	0,9083

Fonte: IBGE (2002), INPI (2001), ISI (2000), IPEA (2003) elaboração própria.

TABELA IV-C

**Estatística das Variáveis contínuas com algum tipo de produção científica e tecnológica
(Universo total / N=562)**

Descrição da Variável	Média	Desvio Padrão	Mínimo	Máximo
Biblioteca	2,0552	5,3370	0,0000	80,0000
Jornal (diário)	6,1352	20,4489	0,0000	480,0000
Tx rede de esgoto	0,8232	0,1542	0,1551	0,9995
Banheiro	0,9772	0,0563	0,4242	1,0000
Coleta de lixo	0,8667	0,1362	0,1598	1,0000
Área urbana	0,8608	0,1520	0,1862	1,0000
Alfabetizado	91,7774	5,1177	54,4000	98,2000
IDH-M	0,7937	0,043	0,5212	0,9119

Fonte: IBGE (2002), INPI (2001), ISI (2000), IPEA (2003) elaboração própria.

Comparando os municípios que possuem registros de patentes em contraposição aos municípios com artigos publicados, verificamos que três das variáveis de infra-estrutura cultural diferenciam o grupo de municípios com patentes do grupo de municípios com artigos. São elas presença de provedor de Internet, número de bibliotecas e número de jornais diários. Nos municípios com artigos, a frequência dessas três variáveis é sempre superior: 81% possui provedor de Internet, o número médio de bibliotecas públicas é de 3,4 e o número médio de jornais diários é 8. (Ver Tabelas V, V-A, V-B e V-C)

TABELA V
Estatística Descritiva para os municípios com patentes (N=494)

Descrição da Variável	Frequência		Percentual	
	Não tem	Tem	Não tem	Tem
Favelas	221	273	44,74	55,26
Incentivos p/ atração de ativ econômica	54	440	10,93	89,07
Programa de Trabalho e Renda	139	355	28,14	71,86
Progr. De capacitação profissional	138	356	27,94	72,06
Provedor de Internet	173	321	35,02	64,98
Livraria	94	400	19,03	80,97

Fonte: IBGE (2002), INPI (2001), elaboração própria.

TABELA V-A
Estatística Descritiva para os municípios com artigos (N=204)

Descrição da variável	Frequência		Percentual	
	Não tem	Tem	Não tem	Tem
Favelas	77	127	37,75	62,25
Incentivos p/ atração de ativ econômica	29	175	14,22	85,78
Programa de Trabalho e Renda	49	155	24,02	75,98
Progr. De capacitação profissional	45	159	22,06	77,94
Provedor de Internet	38	166	18,63	81,37
Livraria	24	180	11,76	88,24

Fonte: IBGE (2002), ISI (1999), elaboração própria.

TABELA V-B
Estatística das Variáveis contínuas para municípios com patente
(Universo total / N=494)

Descrição da Variável	Média	Desvio Padrão	Mínimo	Máximo
Biblioteca	2,1194	5,6100	0,0000	80,0000
Jornal (diário)	5,3462	3,7541	0,0000	27,0000
Tx rede de esgoto	0,8307	0,1493	0,1551	0,9995
Banheiro	0,9826	0,0450	0,4242	1,0000
Coleta de lixo	0,8776	0,1261	0,2997	1,0000
Área urbana	0,8678	0,1470	0,1905	1,0000
Alfabetizado	92,4539	3,9149	54,4000	98,2000
IDH-M	0,7987	0,0374	0,5628	0,9119

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

TABELA V-C
Estatística das Variáveis contínuas para municípios com artigo
(Universo total / N=204)

Descrição da Variável	Média	Desvio Padrão	Mínimo	Máximo
Biblioteca	3,3971	8,5220	0,0000	80,0000
Jornal (diário)	8,0637	33,6222	0,0000	480,0000
Tx rede de esgoto	0,8519	0,1447	0,2937	0,9995
Banheiro	0,9696	0,0646	0,4781	0,9995
Coleta de lixo	0,8862	0,1347	0,1598	1,0000
Área urbana	0,8936	0,1352	0,1862	1,0000
Alfabetizado	91,3196	6,3745	56,0000	97,4000
IDH-M	0,794	0,0515	0,5212	0,919

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

4.2. Resultados do modelo

O modelo econométrico é estimado em duas especificações distintas. Na primeira especificação, utilizamos como variável *proxy* para o desenvolvimento humano o próprio IDH-M e incluímos as demais variáveis independentes que buscam captar infra-estrutura cultural, grau de acesso à informação e existência de programas de incentivo econômico no município. A segunda especificação utiliza os fatores estimados através da análise de componentes principais como *proxies* para o nível de desenvolvimento humano. Lembramos que os dois fatores são estimados utilizando as variáveis: taxa de domicílios em área urbana, taxa de domicílios com coleta de lixo e abastecimento de água, mortalidade infantil e taxa de alfabetização. Existem pelo menos duas vantagens em procedermos às duas estimações: em primeiro lugar, a estimação utilizando o IDH-M permite uma interpretação direta já que o mesmo corresponde a um índice de hierarquização dos municípios quanto ao nível de desenvolvimento humano. Assim, é possível através desta estimação mensurar qual o impacto de uma mudança no nível de desenvolvimento humano sobre a possibilidade de produção tecnológica. Além disso, esta é uma medida universal.

Por outro lado, a dimensão de saúde captada na segunda estimação, utilizando os fatores, é mensurada por meio da taxa de mortalidade infantil e condições de saneamento básico e urbanização, constituindo-se em um indicador mais sensível às mudanças. No IDH-M, a saúde é retratada pela esperança de vida ao nascer que, por ser um indicador de descrição das condições de saúde de todos os grupos etários, apresenta um componente de inércia, estando sujeito a menor variância no tempo. Assim, como a esperança de vida ao nascer é uma média da saúde de toda a população, podem ocorrer mudanças que se compensem e termos municípios com um mesmo valor para esse indicador com condições de saúde bastante diferenciadas entre os grupos etários. Nesse sentido, os dados do IDH-M não refletem completamente a realidade atual e, portanto, a segunda especificação nos permitir inferir o grau de robustez das relações estimadas, na medida em que alteramos os indicadores utilizados como *proxies* para o desenvolvimento humano.

4.2.1. Modelo com IDH-M

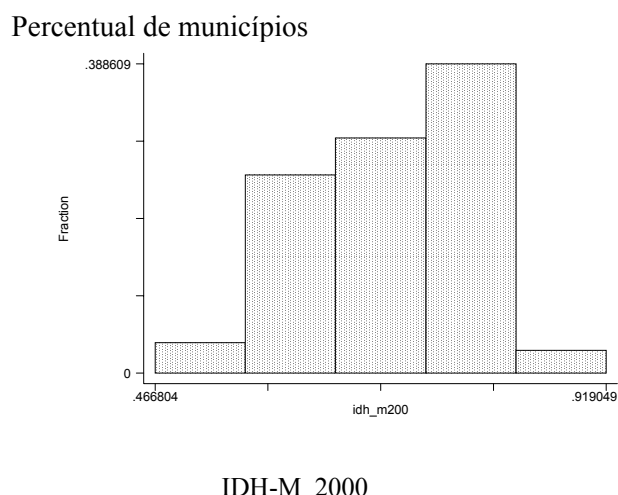
Conforme já dito anteriormente, o Índice Municipal de Desenvolvimento Humano (IDH-M) assume valores entre 0 e 1, constituindo-se em uma variável contínua. No caso dos municípios brasileiros, o IDH-M assume valores entre 0,4668 e 0,919 e a distribuição entre os municípios é assimétrica para a direita (ver Histograma 1). Em função dessas características e a fim de tornar a estimação do modelo logístico possível e a interpretação dos coeficientes factível¹³, construímos variáveis *dummy* que refletem níveis de desenvolvimento humano diferentes entre os municípios. São definidas oito variáveis *dummy* de modo a ter um número de observações relativamente homogêneo¹⁴.

¹³ Como os intervalos entre os IDH-M são infinitesimais, ao utilizarmos a variável em sua forma contínua, o coeficiente deve assumir valores muito elevados para a variação de uma unidade na nossa variável dependente.

¹⁴ As variáveis *dummy* são: *didh1* igual a 1 nos municípios onde o IDH-M assume um valor inferior a 0,60 e *didh1* é igual a 0 para os demais (663 observações); *didh2* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,60 e 0,65 e *didh2* igual a 0 para os demais (779 observações); *didh3* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,65 e 0,70 e *didh3* igual a 0 para os demais (779 observações); *didh4* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,70 e 0,725 e *didh4* igual a 0 para os demais (441 observações); *didh5* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,725 e 0,750 e *didh5* é igual a 0 para os demais (598 observações); *didh6* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,750 e 0,775 e *didh6* igual a 0 para os demais (648 observações); *didh7* é igual a 1 nos municípios onde o IDH-M assume um valor entre 0,775 e 0,80 e *didh7* igual a 0 para os demais (585 observações) e *didh8* é igual a 1 nos municípios com IDH-M superior a 0,80 e igual a zero nos demais (540 observações).

Desse modo, na primeira etapa da estimação do modelo, que se constitui em um modelo logístico, a variável IDH-M é especificada através dessas *dummies*, sendo os municípios de menor IDH-M os municípios de referência, ou seja, aqueles cujo valor do IDH-M é inferior a 0,6.

HISTOGRAMA I
Distribuição de municípios segundo intervalos de IDH-M em 2000



Os resultados encontrados para a primeira etapa do modelo, logit, que estima a decisão de ter ou não uma patente no município revelam que há uma relação positiva entre desenvolvimento econômico e probabilidade de ocorrência de patentes¹⁵. Chama a atenção o elevado grau de ajuste do modelo que apresenta pseudo R^2 igual a 0,4669. Os resultados encontrados mostram que apenas para municípios com o IDH-M superior a 0,725 é que a probabilidade de ter patentes registradas difere dos municípios com pior nível de desenvolvimento. Nesse caso, a probabilidade de ter pelo menos uma patente registrada no município se eleva em cinco vezes. Essa probabilidade é monotonamente crescente à medida que aumenta o nível de desenvolvimento, chegando a ser 78 vezes maior para os municípios com IDH-M maior que 0,80.

¹⁵ Os resultados apresentados na Tabela 5 se referem às razões de chance estimadas e não aos coeficientes. As razões de chance mostram a variação da categoria analisada em relação à categoria de referência. Razões de chance superiores a 1 mostram que os municípios que possuem a categoria analisada apresentam chance superior de apresentar determinada característica em relação aos municípios que não possuem. Por exemplo, observando a tabela 5, verificamos que a razão de chance estimada é de 2,59. Esse valor significa que um município que possui favela ou assemblado tem 159% de chance a mais de ter uma patente registrada. Por outro lado, uma razão de chance inferior a 1, indica que o município tem chance inferior. Uma razão de chance de 0,40, por exemplo, significa que o município tem 60% de chance a menos de ter uma patente registrada.

TABELA VI
Modelo de patentes – Logístico¹⁶

Descrição das Variáveis	Razão de chance	T
Favelas	2.8524	7.88
Incentivos p/ atração de ativ. Econômica	1.9888	3.81
Proger	0.8942	-0.73
Programa de capacitação profissional	1.2851	1.68
Provedor de Internet	2,3845	6.18
Presença de Livrarias	1.5381	2.84
Número de Bibliotecas	1.046	1.19
Número de Jornais (diário)	1.00	0.76
Ano de instalação do município	0.9687	-8.64
Didh2	0.1918	-1.43
Didh3	0.5220	-0.79
Didh4	0.9219	-0.10
Didh5	6.2698	2.93
Didh6	12.9976	4.23
Didh7	22.6645	5.17
Didh8	78.2168	7.17
Pseudo R2		0.4669

Fonte: IBGE (2002), INPI (2001), ISI (2000), IPEA (2002), elaboração própria.

Ademais, a presença de incentivos econômicos, de provedor de Internet e de livrarias contribui para a criação de patentes. No caso de provedor de Internet a probabilidade aumenta em 138% em relação aos municípios que não têm acesso a esse tipo de recurso enquanto a presença de livrarias aumenta em 50% a chance do município ter uma patente registrada. Embora a existência de favelas venha a ser uma característica negativa do desenvolvimento econômico, é também um indicador de grau de urbanização do município. Desse modo, a correlação positiva encontrada pode ser interpretada não porque favela induza a geração de patentes, mas sim como indicativo de quanto maior o centro urbano maior a probabilidade de existência de patentes. Por outro lado, a não significância estatística da variável número de bibliotecas públicas chama atenção, porque espera-se que seja uma variável importante (vide Tabela VI).

¹⁶ A categoria de referência é um município com IDH-M de valor inferior a 0,60, sem favelas, sem livrarias, sem programa de capacitação profissional, sem programas de incentivo e renda e sem provedor de internet.

TABELA VII
Modelo de artigos – Logístico¹⁷

Descrição das Variáveis	Razão de chance	T
Favelas	2.0928	4.20
Incentivos p/ atração de ativ. Econômica	1.1045	0.42
Proger	0.9590	-0.20
Programa de capacitação profissional	1,3199	1.37
Provedor de Internet	5.607	8.00
Presença de Livrarias	2.1795	3.04
Número de Bibliotecas	1.2088	3.86
Número de Jornais (diário)	1.007	1.93
Ano de instalação do município	0.9848	-2.47
Didh2	0.5612	-0.57
Didh3	1.4524	0.46
Didh4	1.5157	0.49
Didh5	2.5917	1.23
Didh6	3.9690	1.86
Didh7	4.0578	1.88
Didh8	9.8131	3.14
Pseudo R2		0.3807

Fonte: IBGE (2002), INPI (2001), ISI (2000), IPEA (2002), elaboração própria.

No que tange à publicação de artigos (Tabela VII), os resultados são bastante similares, com exceção das variáveis presença de incentivos às atividades econômicas, número de bibliotecas públicas e número de jornais diários. Nesse caso, o número de bibliotecas e o número de jornais diários são estatisticamente significativos, sugerindo uma diferença no processo de produção de artigos em relação ao de patentes. No caso de artigos, provavelmente, esse resultado se deve à necessidade de densidade de massa crítica, algo comum a localidades onde estejam presentes instituições de pesquisas, tais como universidades. Além disso, a magnitude do coeficiente estimado para presença de livrarias no modelo de publicação de artigos é superior ao estimado no modelo de registro de patentes, o que pode reforçar a interpretação sobre o papel que as instituições de pesquisa assumem. Quanto às variáveis do IDH-M, constatamos que somente a partir da *dummy* Didh6, os coeficientes passam a ser estatisticamente significativos¹⁸, sugerindo que apenas os municípios onde o índice de desenvolvimento apresenta valor superior a 0,750 têm chances de publicar artigos.

Na segunda etapa do modelo de registro de patentes, estimação do modelo Binomial Negativo, não utilizamos as *dummies* que retratam a escala de IDH-M entre os municípios, mas sim o próprio índice¹⁹. Essa especificação da variável se deve a dois motivos: em primeiro lugar a especificação do IDH-M através de variáveis *dummy* aumenta em muito o número de variáveis independentes. Segundo, a especificação da variável na forma contínua, no caso do modelo binomial negativo admite

¹⁷ A categoria de referência é um município com IDH-M de valor inferior ou igual a 0,35, sem favelas, sem livrarias, sem programa de capacitação profissional, sem programas de incentivo e renda e sem provedor de internet.

¹⁸ Os coeficientes são significativos a 6%.

¹⁹ O modelo utilizando a especificação das *dummies* de IDH-M também foi rodado, mas os coeficientes não são significativos. Este resultado deve-se, provavelmente, aos intervalos escolhidos. Como o número de municípios com registro positivo de artigos e patentes é muito pequeno, optamos por apresentar apenas a especificação do IDH-M na forma contínua.

interpretação direta. Os resultados apontam para significância estatística das variáveis presença de favelas, presença de provedor de Internet, número de bibliotecas e para o IDH-M. Nesse caso, reportamos os coeficientes estimados na Tabela VIII. Para as variáveis binárias, calculamos o *antilog*²⁰ em termos percentuais e, para as variáveis contínuas ou discretas, o coeficiente pode ser interpretado diretamente como uma semi-elasticidade (Cameron e Trivedi, 1988). Os resultados corroboram os obtidos na primeira etapa, para aqueles municípios que possuem favela, o número esperado de patentes registradas é 345% maior em comparação aos que não possuem favelas. Quanto a Internet, municípios com provedor de Internet apresentam um número esperado de patentes 206% maior do que o obtido em municípios que não possuem. Chama atenção o resultado obtido para a variável número de bibliotecas que, na primeira etapa, ou seja, na decisão de registrar patentes, não se mostra significativo e para definir o número de patentes, torna-se significativo. A presença de uma biblioteca pública adicional implica que o número esperado de patentes aumenta em 20%. Além disso, uma melhora no índice de desenvolvimento humano em uma unidade tende a ampliar em 1000% o número esperado de patentes²¹.

TABELA VIII
Modelo de patentes: Binomial Negativo truncado ao zero²²

Descrição das Variáveis	Coeficientes	t
Favelas	1,4939	6,62
Incentivos p/ atração de ativ. Econômica	-0,1583	-0,45
Proger	-0,0505	-0,02
Programa de capacitação profissional	0,4964	1,85
Provedor de Internet	1,1199	4,15
Livraria	0,2450	0,81
Biblioteca	0,2092	2,68
Jornal (diário)	0,0263	0,92
Ano de instalação do município	-0,0117	-1,26
IDH-M	10,3461	3,81
Constante	-17,2315	-1,30
Pseudo R2	0,0697	

Fonte: IBGE (2002), INPI (2001), ISI (2000), IPEA (2002), elaboração própria.

No modelo Binomial Negativo para artigos, apenas as variáveis número de bibliotecas públicas, presença de livrarias e IDH-M são significativos a 5%. Embora o nível de significância não seja o mesmo do modelo *logit* (primeira etapa), as variáveis de infra-estrutura cultural permanecem como determinantes do número esperado de artigos publicados. A presença de mais uma biblioteca no município gera um número esperado de artigos 16% maior. (Tabela IX).

²⁰ O “antilog” é dado pelo $(\exp \beta - 1) \cdot 100$.

²¹ O IDH-M varia no intervalo de 0,4668 a 0,9190 com média de 0,70 e desvio padrão de 0,081. Não existe uma interpretação direta para o significado da variação de uma unidade no IDH-M. No caso do modelo com as *dummies*, o desenvolvimento humano não é estatisticamente significativo na segunda etapa dos modelos.

²² O modelo converge após 14 iterações, entretanto a função de máxima verossimilhança não é côncava e desse modo não podemos inferir que a função converge em um ponto de máximo global.

TABELA IX
Modelo de artigos: Binomial Negativo truncado ao zero²³

Descrição das Variáveis	Coefficientes	T
Favelas	0,3641	0,81
Incentivos p/ atração de ativ. Econômica	-0,6424	-0,85
Proger	0,4626	0,55
Programa de capacitação profissional	0,5747	0,65
Provedor de Internet	-1,2550	-0,82
Livraria	1,6398	2,05
Biblioteca	0,1646	2,82
Jornal (diário)	-0,0104	-1,29
Ano de instalação do município	-0,01932	-1,09
IDH-M	12,7316	2,87
Constante	-16,1470	-6,56
Pseudo R2	0,0496	

Fonte: IBGE (2002), INPI (2001), ISI (2000), FJP (1997), elaboração própria.

4.2.2. Modelo com indicador de desenvolvimento urbano (f1) e indicador de presença de pobreza (f2)

No caso da segunda especificação, em que incluímos os fatores f1 e f2 como covariadas e retiramos o IDH-M, os resultados encontrados para a primeira etapa do modelo, logit, surpreendem pela robustez do conjunto de variáveis, na medida em que das onze covariadas apenas quatro não são significativas estatisticamente e, sobretudo, os sinais são os esperados. O indicador de desenvolvimento urbano (f1), resultante do método dos componentes principais, indica uma probabilidade maior para os municípios mais desenvolvidos, ao passo que o indicador de presença de pobreza (f2) mostra que os municípios com maior incidência de pobreza têm uma chance menor de registrar patentes, uma vez que a razão de chance estimada é menor do que 1.²⁴ No que tange às demais variáveis independentes, os resultados são similares à especificação com IDH-M (Tabela X).

²³ O modelo converge após 24 interações, entretanto a função de máxima verossimelhança não é côncava e desse modo não podemos inferir que a função converge em um ponto de máximo global.

²⁴ Como os indicadores f1 e f2 são variáveis contínuas, a interpretação da razão de chances não é a mesma de uma variável binária. Entretanto, uma razão de chance menor que 1 pode ser interpretada como uma redução na probabilidade.

TABELA X
Modelo de patentes – Logístico²⁵

Descrição das Variáveis	Razão de chance	T
Favelas	2,366886	6,66
Incentivos p/ atração de ativ. econômica	2,014382	3,92
Programa de Trabalho e Renda	1,000166	0,00
Programa de capacitação profissional	1,304687	1,80
Provedor de Internet	1,908577	4,57
Presença de Livrarias	1,697474	3,54
Número de Bibliotecas	1,069584	1,35
Número de Jornais (diário)	0,9976649	-0,73
Ano de instalação do município	0,9834522	-4,31
F1	5,481485	10,35
F2	0,5647159	-3,32
Pseudo R2		0,4869

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

Na Tabela XI, apresentamos os resultados para o modelo logit referente à publicação de artigos, os resultados são também bastante similares, com exceção das variáveis presença de incentivos às atividades econômicas, número de bibliotecas públicas e número de jornais diários quando comparado ao de registro de patentes da segunda especificação e com exceção de número de jornais diários, quando comparado com o de publicação de artigos da primeira especificação. O resultado surpreendente é a magnitude da razão de chance estimada para o indicador f2, maior do que 1, o que pode ser interpretado como quanto maior a incidência de pobreza no município, maior a chance de publicação de artigos. Como o indicador f2 capta a incidência de pobreza, ele atribui valores positivos tanto para cidades homogeneamente pobres como para cidades com elevados níveis de desigualdade (como os grandes centros urbanos do país). Municípios como São Paulo, Campinas, Rio de Janeiro e Belo Horizonte (que apresentam o indicador f2 maior do que 1) ilustram localidades com expressiva produção científica e com bolsões de pobreza.

TABELA XI
Modelo de artigos – Logístico

Descrição das Variáveis	Razão de chance	T
Favelas	1,701835	3,09
Incentivos p/ atração de ativ. econômica	1,16937	0,66
Programa de Trabalho e Renda	0,9918078	-0,04
Programa de capacitação profissional	1,221111	1,00
Provedor de Internet	4,26396	6,92
Livraria	2,028063	2,75
Biblioteca	1,211695	3,78
Jornal (diário)	1,006608	2,65
Ano de instalação do município	0,9949102	-0,75
f1	2,147485	6,82
f2	1,723003	4,11
Pseudo R2		0,4043

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

²⁵ As categoria de referência é um município sem favelas, sem livrarias, sem programa de capacitação profissional, sem programas de incentivo e renda e sem provedor de internet.

As Tabelas XII e XIII reportam os resultados do modelo de registro de patentes e publicação de artigos de acordo com a estimação do Binomial Negativo. Não há divergência da robustez das variáveis com relação à especificação com IDH-M. Nesse caso, o grau de desenvolvimento urbano (f1) e o indicador de presença de pobreza (f2) são estatisticamente significativos.

TABELA XII
Modelo de patentes: Binomial Negativo truncado ao zero

Descrição das Variáveis	Coefficientes	T
Favelas	0,9920797	4,71
Incentivos p/ atração de ativ. econômica	-0,0090898	-0,03
Programa de Trabalho e Renda	0,0565973	0,22
Programa de capacitação profissional	0,5674528	2,38
Provedor de Internet	0,8584361	3,19
Livraria	0,3221403	1,10
Biblioteca	0,1317063	2,54
Jornal (diário)	0,0130499	0,47
Ano de instalação do município	-0,0027255	-0,33
f1	0,7145721	5,08
f2	0,4008395	2,17
Constante	-3,1470520	-3,10

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

No que tange à estimativa do número de artigos publicados, mais uma vez somente as variáveis número de bibliotecas públicas e as referentes ao desenvolvimento do município, indicadores de grau de desenvolvimento urbano e de nível de pobreza são significativos a 5%. Presença de livrarias e número de jornais diários são significativos a 12% e 11%, respectivamente. Desse modo, a conclusão obtida para a primeira especificação (produção de artigos com IDH-M) com o mesmo processo de estimação (Binomial Negativo) vale para a especificação com f1 e f2, ou seja, as variáveis de infra-estrutura cultural permanecem como determinantes do número esperado de artigos publicados. A presença de mais uma biblioteca no município gera um número esperado de artigos 17% maior.

TABELA XIII
Modelo de artigos: Binomial Negativo truncado ao zero

Descrição das Variáveis	Coefficientes	t
Favelas	0,3758748	0,84
Incentivos p/ atração de ativ. Econômica	-0,5287083	-0,69
Programa de Trabalho e Renda	0,6159782	0,92
Programa de capacitação profissional	-0,1645596	-0,17
Provedor de Internet	-1,5433260	-1,15
Livraria	1,3297150	1,53
Biblioteca	0,1741710	3,14
Jornal (diário)	-0,0115120	-1,58
Ano de instalação do município	-0,0215057	-1,25
F1	0,6739465	3,21
F2	1,7192260	3,53
Constante	-9,4287250	-7,55

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

Na comparação entre as duas etapas da estimação e as especificações, observamos que o ajuste do modelo *logit* é melhor que do Binomial Negativo, uma vez que o pseudo R^2 do primeiro é bastante elevado, estando em torno de 0,4 em ambas especificações, ao passo que no Binomial Negativo assume valores próximos a 0,06. Possivelmente, o reduzido número de observações no modelo Binomial Negativo (494 municípios com patentes e 204 municípios com artigos) afete a estimação do mesmo.²⁶

Realizamos o teste de especificação do modelo *Hurdle* contra o modelo de Poisson para as duas situações (registro de patentes e produção de artigos) e em ambas especificações. Nesses quatro casos, o modelo de Poisson é rejeitado, ratificando a utilização do *Hurdle*.

Os resultados obtidos no modelo *Hurdle* podem ser sintetizados conforme nossos grupos de variáveis. A infra-estrutura cultural (presença de provedor de Internet e de livrarias, número de bibliotecas e de jornais diários) explica as decisões de registrar patentes e de publicar artigos. A segunda etapa do processo decisório, quanto de patentes e artigos produzir, estimado pelo modelo Binomial Negativo, mostra que a infra-estrutura cultural já não é tão importante, pois apenas as variáveis número de bibliotecas e de provedor de Internet são determinantes.

As *dummies* para escala do IDH-M dos municípios com maior grau de desenvolvimento humano e o próprio IDH-M assim como os grupos de variáveis de infra-estrutura urbana, de saúde e de educação, redefinidos, por meio do método de componentes principais, como indicador de desenvolvimento urbano (f1) e como a presença de pobreza do município (f2) são determinantes das duas etapas do processo decisório referente à produção científica e tecnológica. Em outras palavras, quanto mais desenvolvido e quanto maior a presença de pobreza, maior a chance de registrar patentes e publicar artigos e maior o número esperado de patentes como também de artigos.

A variável presença de favelas ou assemelhados²⁷ no município está, também, correlacionada com a decisão de produzir patentes/artigos e a decisão quanto ao número de patentes a serem registradas. E sua importância deve ser explicada pelo grau de urbanização do município, porque quanto maior o centro urbano, maior a probabilidade de existência de favelas assim como de condições favoráveis de infra-estrutura cultural e desenvolvimento urbano.

5. CONCLUSÕES PRELIMINARES

Este trabalho é um esforço inicial para uma avaliação mais abrangente das complexas relações entre produção científica e tecnológica e bem-estar social. A contribuição do artigo está na avaliação estatística realizada, tomando o município como nível de análise e utilizando dados disponíveis de ciência, tecnologia e bem-estar social.

Duas conclusões preliminares podem ser mencionadas:

²⁶ No modelo Binomial Negativo, a função de máxima verossimilhança converge, mas não parece ser côncava.

²⁷ A princípio, essa variável integrava o grupo de variáveis referentes à infra-estrutura urbana. No entanto, o método de componentes principais indicou que ela deveria ser excluída desse grupo redefinido em f1 e f2.

- em primeiro lugar, há uma forte relação entre a infra-estrutura cultural e a produção científica e tecnológica;
- em segundo lugar, há uma complexa associação entre a produção científica e tecnológica e os dados de bem-estar social. Os resultados dos testes realizados indicam que a produção científica e tecnológica está associada tanto a indicadores de desenvolvimento humano como a indicadores de presença e incidência de pobreza.

Como compreender essas duas conclusões?

A associação direta e clara entre dados de infra-estrutura cultural (bibliotecas, jornais, internet) e desenvolvimento tecnológico não é de difícil compreensão. Certamente a disponibilidade de mais recursos desse tipo é uma pré-condição para a produção científica e tecnológica.

Já a correlação entre produção científica e tecnológica e desenvolvimento e incidência da pobreza exige uma discussão mais cuidadosa. Pelo menos quatro argumentos podem ser considerados nessa explicação.

Em primeiro lugar, a distribuição espacial das atividades de ciência e tecnologia de certa forma acompanha o mapa brasileiro de concentração regional de renda e de atividades econômicas. Por isso, a obtenção de um IDH correspondente ao de um país de “alto desenvolvimento humano” ($IDH > 0,8$) tem efeitos tão significativos na probabilidade de identificação de produção científica e tecnológica em um município.

Em segundo lugar, essa correlação reflete um elemento estrutural do desenvolvimento brasileiro. Ou seja, embora exista associação entre desenvolvimento tecnológico e desenvolvimento humano, o desenvolvimento humano identificado é concentrado, gerando bolsões de pobreza. As desigualdades não apenas identificam-se no nível regional, mas também no interior das cidades. Talvez essa seja a maior contribuição da combinação entre os indicadores f1 e f2 propostos no texto: uma grande cidade simultaneamente obtém altos valores nos dois indicadores, demonstrando a combinação entre desenvolvimento e presença de pobreza (identificada de outra forma pela associação entre IDH alto e presença de favela). O IDH por si só não indica essa combinação (São Paulo tem $IDH = 0,841$; Campinas $IDH = 0,852$; Rio de Janeiro $IDH = 0,842$; Belo Horizonte $IDH = 0,839$). As principais cidades do país, grandes centros urbanos, comportam enormes desigualdades sociais.²⁸ Inúmeros estudos apontam esse padrão. Um importante estudo que poderia ser recuperado para essa discussão é o estudo do CEBRAP (Camargo, Cardoso et al, 1976): *São Paulo 1975: crescimento e pobreza*. Os dados avaliados nas seções anteriores demonstram a persistência desse padrão perverso de desenvolvimento.

Uma outra forma de interpretação desse resultado diz respeito ao impacto da infra-estrutura científico-tecnológica sobre as condições de vida. A associação encontrada entre infra-estrutura científico-tecnológica e condições de vida aqui identificadas sugere que há uma relação causal na direção entre ciência e tecnologia e desenvolvimento humano, porém essa não é tratada neste trabalho,

²⁸ Segundo dados do IBGE, existem 1,077 milhões de pessoas vivendo em favelas, cortiços, ruas e “buracos” da cidade de São Paulo (correspondendo a 10,3% da população da cidade) (*Folha de São Paulo*, 28/07/2002, p. C6).

podendo ser incorporada em pesquisa futura, permitindo avaliar a endogeneidade entre infra-estrutura científico-tecnológica e desenvolvimento humano.

Sabendo-se que a infra-estrutura científica desenvolveu-se mais do que a tecnológica nas últimas três décadas, por que essa infra-estrutura ainda não consegue gerar condições de bem-estar? Dois fatores podem estar atuando nesse ponto: por um lado, a produção científica brasileira ainda não conseguiu acumular massa crítica suficiente para influenciar de forma perceptível a realidade social; por outro lado, podem existir desconexões entre as linhas de concentração das atividades científicas e as necessidades sociais mais urgentes da população mais pobre do país ou, talvez, uma combinação entre esses elementos de falta de massa crítica e desconexão parcial.

Embora preliminares, essas conclusões podem colocar em discussão algumas sugestões de políticas públicas:

1. investimentos para a expansão da infra-estrutura científico-tecnológica pelo país, de forma a mitigar a sua elevada concentração regional (apenas 204 municípios publicaram ao menos um artigo indexado em 1999 e apenas 494 municípios são sede de firmas/instituições que registraram ao menos uma patente entre 1988 e 1996): uma distribuição menos concentrada pode ser uma forma dessa infra-estrutura dar conta da diversidade de problemas existentes no país, alcançando municípios com IDH mais baixo, por exemplo;
2. fortalecimento da infra-estrutura cultural: um pré-requisito importante para a produção científico-tecnológica e, como demonstrado no trabalho, influi sobre a produção científica e tecnológica;
3. como a teoria sugere e várias evidências são coletadas, avanços sociais têm impacto sobre a produção científica e tecnológica, seja pela melhoria na saúde (com impactos sobre capacidade de aprendizado), redução do analfabetismo e aprimoramento educacional;
4. quanto à infra-estrutura científica e tecnológica, dois outros movimentos podem ser realizados: em primeiro lugar, a ampliação significativa dos recursos à disposição do setor (uma pré-condição para a acumulação de massa crítica necessária para acionar um impacto positivo das atividades científicas e tecnológicas sobre as condições de vida); em segundo lugar, um melhor direcionamento dos recursos existentes no sentido do atendimento de prioridades sociais como saúde, habitação e condições de vida em grandes centros urbanos. Nessa linha, podem ser encaminhados projetos “orientados-por-missão”, como sugerido por Freeman (1996) para condições ambientais, sugestão que poderia ser re-orientada para “missões” de forte impacto social.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMOVITZ, M. (1989) *Thinking about growth*. Cambridge: Cambridge University.
- CAMARGO, C. F. C.; CARDOSO, F. H.; MAZZUCHELLI, F.; MOISÉS, J. A.; et all (1976) *São Paulo 1975: crescimento e pobreza*. São Paulo: Edições Loyola.
- CAMERON, A. C., TRIVEDI, P. K, MILNE, Frank, PIGGOTT, J. A microeconomic model of the demand for health care and health Insurance in Australia. *Review of Economics Studies*. vol. 55, págs. 85-106, 1988.
- FREEMAN, C. (1996) The greening of technology and models of innovation. *Technological Forecast and Social Change*, v. 53, pp. 27-39.
- FOGEL (1994) Economic growth, population theory, and physiology on the making of economic policy. *American Economic Review*, v. 84, n. 3, Jun.
- GRILICHES, Z. (1990) Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, v. 28, Dec.
- GLOBAL FORUM FOR HEALTH RESEARCH (2002) *The 10/90 Report on health research 2001-2002*. Geneva: Global Forum for Health Research.
- MURRAY, C. J. (ed.); LOPEZ, A. (1996) *The global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected*. Harvard: Harvard School of Public Health.
- NORONHA, K. V. M. S. Dois Ensaio sobre a desigualdade social em saúde. Belo Horizonte: CEDEPLAR/UFGM, 105p. Dissertação de mestrado, 2001.
- UNITED NATIONS DEVELOPMENT PROGRAM (2001) *Human Development Report: making new technologies work for human development*. New York: UNDP (www.undp.org)
- WANG, J.; JAMISSON, D.; BOS, E.; PREKER, A.; PEABODY, J. (1999) *Measuring country performance in health: selected indicators for 115 countries*. Washington: World Bank (disponível em www.worldbank.org).
- WOOLDRIDGE, J. M. *Econometric Analysis of Cross-Section and Panel Data*, EMD Press, 2002.
- WORLD BANK (1993) *World Development Report 1993: Investing in Health*. Oxford: Oxford University.
- WORLD HEALTH ORGANIZATION (1996) *Investing in health research and development*. TDR/Gen/96.1 (www.who.org).
- WORLD HEALTH ORGANIZATION (1999) *World Health Report: making a difference*. Geneva: who (www.who.org).
- WORLD HEALTH ORGANIZATION (2001) *Macroeconomics and health: investing in health for economic development*. Report of the Comission on Macroeconomics and Health. Geneva (www.who.org).

ANEXO - APÊNDICE DE TABELAS

Matriz de Correlação das variáveis independentes

Descrição da Variável	Biblioteca	Livraria	Jornal (diário)	Tx mort infantil	Abastec água via rede geral*	Banheiro*	Coleta de lixo*	Alfabetizados	Programa de trabalho e renda	Favela	Internet	Área urbana*	Capacitação Profissional
Biblioteca	1,0000												
Livraria	0,1264	1,0000											
Jornal (diário)	0,0387	0,0684	1,0000										
Tx. mort. infantil	-0,0733	-0,1845	-0,0657	1,0000									
Abastec. água via rede geral*	0,1452	0,2698	0,0892	-0,2761	1,0000								
Banheiro*	0,1026	0,2739	0,0859	-0,6194	0,4790	1,0000							
Coleta de lixo*	0,1628	0,3468	0,1261	-0,3749	0,7451	0,6690	1,0000						
Alfabetizados	0,1190	0,3179	0,0960	-0,8098	0,4025	0,7886	0,5662	1,0000					
Programa de trabalho e renda	0,0892	0,2287	0,0517	-0,0536	0,1157	0,0978	0,1553	0,1279	1,0000				
Favela	0,0940	0,1948	0,0491	0,0446	0,0892	0,0597	0,1182	0,0288	0,1779	1,0000			
Internet	0,1713	0,4107	0,1025	-0,1866	0,3396	0,2578	0,4075	0,3166	0,2117	0,2111	1,0000		
Área urbana*	0,1618	0,3290	0,1054	-0,2366	0,7507	0,5007	0,8424	0,4119	0,1244	0,1544	0,4082	1,0000	
Capacitação Profissional	0,0982	0,2189	0,0639	-0,0098	0,1087	0,0724	0,1494	0,0890	0,4687	0,1692	0,2138	0,1453	1,0000

Fonte: IBGE (2002), elaboração própria.

* A variável se refere ao percentual de domicílios atendidos (ou situados, no caso de área urbana).

Matriz de Correlação das variáveis independentes - dummy produção intelectual = 1

Descrição da Variável	Biblioteca	Livraria	Jornal (diário)	Tx mort infantil	Abastec água via rede geral*	Banheiro*	Coleta de lixo*	Alfabetizados	Programa de trabalho e renda	Favela	Internet	Área urbana*	Capacitação Profissional
Biblioteca	1,0000												
Livraria	0,0915	1,0000											
Jornal (diário)	0,0168	-0,0728	1,0000										
Tx mort infantil	-0,0605	-0,1030	-0,0268	1,0000									
Abastec água via rede geral*	0,1297	0,1778	0,0710	-0,1605	1,0000								
Banheiro*	0,0327	0,1314	0,0028	-0,6193	0,3937	1,0000							
Coleta de lixo*	0,1305	0,1930	0,0454	-0,3368	0,6820	0,5719	1,0000						
Alfabetizados	0,1062	0,2152	-0,0072	-0,7171	0,3322	0,8102	0,5309	1,0000					
Programa de trabalho e renda	0,1013	0,1928	-0,0444	-0,0225	0,0570	0,0383	0,0809	0,1269	1,0000				
Favela	0,1274	0,1439	0,0652	0,0960	0,0951	0,0075	0,1408	0,0680	0,1941	1,0000			
Internet	0,1333	0,3616	-0,0140	-0,0783	0,3618	0,1954	0,3645	0,2275	0,2040	0,1467	1,0000		
Área urbana*	0,1314	0,2359	0,0571	-0,1320	0,6950	0,3850	0,8384	0,3541	0,1076	0,2189	0,3911	1,0000	
Capacitação Profissional	0,1123	0,1822	0,0468	0,0041	0,0971	0,0791	0,1324	0,1460	0,4797	0,1309	0,2584	0,1570	1,0000

Fonte: IBGE (2002), INPI (2001), ISI (2000), elaboração própria.

*A variável se refere ao percentual de domicílios atendidos (ou situados, no caso de área urbana)