UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

UMA ANÁLISE EMPÍRICA DO CONSUMO DE MÚSICA NA UNIVERSIDADE DE
SÃO PAULO

Igor Siqueira Cortez

Orientador: Prof. Dr. Gabriel de Abreu Madeira

SÃO PAULO

2010

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Carlos Roberto Azzoni
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Denisard Cneio de Oliveira Alves
Chefe do Departamento de Economia

Prof. Dr. Dante Mendes Aldrighi
Coordenador do Programa de Pós-Graduação em Economia

IGOR SIQUEIRA CORTEZ

**UMA ANÁLISE EMPÍRICA DO CONSUMO DE MÚSICA NA UNIVERSIDADE DE
SÃO PAULO**

Dissertação apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para a obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Gabriel de Abreu Madeira

SÃO PAULO

2010

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Cortez, Igor Siqueira

Uma análise empírica do consumo de música na Universidade de São Paulo / Igor Siqueira Cortez. – São Paulo, 2010.
68 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2010.
Orientador: Gabriel de Abreu Madeira.

1. Direito de propriedade 2. Música 3. Internet 4. Pirataria 5.
Econometria I. Universidade de São Paulo. Faculdade de Economia,
Administração e Contabilidade II. Título.

CDD – 342.123

Aos meus Pais.

Meus primeiros agradecimentos são dedicados aos meus pais pelo apoio e conforto incondicional propiciado ao longo desses 25 anos. Sem essa força eu certamente não teria sucesso na busca dos meus objetivos. Agradeço também a minha irmã que sempre foi uma grande companheira, a minha Vó, Dona Waldellice, que também sempre me apoiou e foi fundamental para que eu conseguisse prosseguir com meus estudos na Universidade de São Paulo.

É impossível não se lembrar dos amigos que tornaram esta jornada possível e muito mais agradável. Entre eles agradeço ao meu grande amigo Henrique (Sensa), Raul e à Tainá que na biblioteca do TST compartilharam conhecimentos e força de vontade para que conseguíssemos nos preparar para o exame da Anpec.

Agradeço imensamente aos amigos que fiz durante o mestrado (Alex, Raphão, Bruno Gasperini, Danigri, Rubens-Hanz, Gustavão, Tinoco, Eduardo Jardim, Vivi, Carla, Mirela, Raquel, James Dean, Joço, Sabba, Malanzinho, Silvio, Genta, Dankas, Attuy e Paula) que propiciaram inúmeros momentos de descontração, mas que também contribuíram imensamente na minha formação acadêmica.

Agradeço ao Leandro, Bruno e Eduardo pela convivência de quase dois anos na república do Apiritiviz e pelas pizzas que dividimos durante esse tempo.

Entre os professores agradeço imensamente ao Prof. Dante que tornou viável o meu projeto de dissertação, ao Prof. Denisard que tornou possível a construção da base de dados que é o meu objeto de estudo. Também gostaria de agradecer pelas contribuições dos professores que ministraram as matérias que cursei (Lucia Barroso, Chiappin, Gabriel Madeira, Fernando Botelho, Mauro, Danilo, Vera , Naércio e Postali). Dentre esses agradeço infinitamente ao meu orientador, Gabriel Madeira, que, sempre com ótimas idéias e comentários pertinentes, participou ativamente e com grande entusiasmo no desenvolvimento dessa dissertação. Agradeço também aos professores Mauro Rodrigues, Marcos Rangel, Fernando Botelho e Eduardo Haddad que contribuíram com ótimos comentários durante a mini-banca, seminário de workshop e qualificação.

Por fim, gostaria de agradecer ao apoio financeiro da FAPESP que permitiu a minha estadia e o desenvolvimento de meus estudos em São Paulo.

RESUMO

O compartilhamento de músicas pela Internet pode ter efeitos negativos ou positivos sobre a indústria musical. Por um lado, pode limitar a poder de incentivos à criação garantidos por direitos autorais. Por outro lado, contribui para a divulgação de artistas e músicas. Este artigo estima os efeitos de *downloads* de música sobre o consumo de CDs e shows para uma amostra de 7.147 estudantes universitários. Utilizando variáveis instrumentais motivadas por um modelo teórico simples, associadas à habilidade em informática e tolerância à ilegalidade, estimamos que a prática de *downloads* de música reduz a probabilidade de consumir CDs em até 45%, ao passo que aumenta a probabilidade de ir a shows em até 35%. Nossos resultados também apontam que existe uma relação de substituição entre pirataria física de CDs e *downloads* de música. Esse último resultado sugere que as estimativas feitas sem levar em conta o consumo de CDs piratas na equação estrutural podem ser viesadas.

ABSTRACT

Music file sharing may have positive or negative effects over the music industry. On the one hand, it might undermine incentives for music creation guaranteed by copyrights. On the other hand, it contributes to the dissemination of artists and songs. In this paper we estimate the effect of music downloads over the consumption of CDs and concerts for a sample of 7.147 university students. Using theoretically motivated instrumental variables related to computer skills and illegality acceptance, we estimate that the practice of downloads generates a reduction on the probability of buying CDs of up to 45% and an increase in the order of 35% on the probability of going to concerts. We also show the existence of a substitution relationship between physical CD piracy and music downloads. This last result suggests that estimates without controlling for CD piracy may be biased.

SUMÁRIO

LISTA DE TABELAS	2
1 INTRODUÇÃO	3
2 LITERATURA.....	7
3 FUNDAMENTAÇÃO TEÓRICA	11
4 DADOS E ESTATÍSTICAS DESCRITIVAS	16
5 ESTRATÉGIA DE IDENTIFICAÇÃO	21
5.1 Resultados de Primeiro Estágio.....	27
6 ABORDAGEM EMPÍRICA	30
7 ANÁLISE DE ROBUSTEZ.....	37
7.1 Mudando Especificações.....	37
7.2 Análise de Subamostras	41
8 DISCUSSÃO E CONCLUSÃO	44
REFERÊNCIAS	46
APÊNDICES	48

LISTA DE TABELAS

Tabela 1 - Estatísticas Sumárias	18
Tabela 2 - Testes de Komogorolv - Smirnov	19
Tabela 3 - Estatísticas Descritivas dos Instrumentos	23
Tabela 4 - Relação dos Instrumentos com Interesse Declarado por Música - Mínimos Quadrados Ordinários	25
Tabela 5 - Estimativas de Primeiro Estágio - Mínimos Quadrados Ordinários	29
Tabela 6 - Estimativas de MQO e Variáveis Instrumentais - Probabilidade de comprar CDs e ir a Shows	31
Tabela 7 - Efeito dos <i>Downloads</i> e CDs Piratas Sobre o Gasto com Shows - GMM.....	34
Tabela 8 - Efeitos de <i>Downloads</i> sobre Quantidade de CDs - TOBIT	34
Tabela 9 - Estimativas de Primeiro Estágio – Análise de Robustez - MQO.....	37
Tabela 10 - Análise de Robustês - Equação Estrutural - GMM.....	38
Tabela 11 - Alterando a Equação Estrutural - GMM.....	40
Tabela 12 - Efeito dos <i>Downloads</i> sobre Shows e CDs – Análise de Subamostras - GMM ..	42

1.INTRODUÇÃO

O advento recente da prática de compartilhamento de músicas pela Internet originou um intenso debate na indústria de entretenimento e no meio acadêmico. Desde o surgimento do *Napster*¹, em 1999, houve uma notável expansão da oferta de músicas disponíveis para *download*² e, simultaneamente, observou-se um declínio nas vendas de CDs de música em todo o mundo. Este fato gerou preocupações sobre possíveis problemas típicos de bens públicos na indústria de música. Ao deixarem de pagar os direitos autorais implícitos no preço de CDs, indivíduos estariam removendo incentivos necessários para a produção de música.

Do ponto de vista teórico, no entanto, não é completamente claro que a disponibilidade de *downloads* tenha um efeito deletério sobre a produção de música. A disponibilidade de músicas na Internet pode a princípio contribuir para a divulgação de artistas e difusão do seu trabalho³. Isso poderia contribuir para o consumo de bens associados à música, entre os quais shows e CDs. Esta indeterminação sobre os efeitos do advento dos *downloads* sobre o mercado de música motivou uma importante literatura empírica. Embora nem sempre os estudos sejam conclusivos (e.g. Oberholzer-Gee & Strumpf (2007)) alguns trabalhos como o de Rob&Waldfoegel(2006) e Zentner(2006) encontram evidências de que *downloads* geram um efeito negativo sobre o consumo de CDs.

Tal literatura baseia-se fundamentalmente em dados sobre consumo de CDs em países desenvolvidos. No entanto, há motivos para se imaginar que o grau de desenvolvimento das instituições legais seja um ingrediente importante na relação entre disponibilidade de músicas para *download* na Internet e compra de CDs. Quando comparamos a evolução das receitas da indústria fonográfica mundial com a brasileira (Gráfico 1), observa-se que no Brasil esta queda ocorreu de forma mais intensa. Este comportamento diferenciado pode estar relacionado à dificuldade de se coibir o consumo ilegal de músicas na Internet⁴ e à forte

¹ Programa pioneiro que permitia a troca de músicas entre seus usuários pela internet. A partir de programas como esse, as pessoas podiam baixar e ouvir as músicas livremente sem ter que pagar pelo direito autoral.

² Ato de baixar arquivo pela internet.

³ Ver Curien et al (2004).

⁴ No Brasil ainda não existe uma lei que prevê expressamente uma punição para pessoas que baixem conteúdo protegido pela Internet.

presença do comércio de produtos de origem ilegal⁵. Os trabalhos empíricos já realizados na tentativa de explicar a queda no consumo de CDs deixam uma lacuna ao ignorar a influência da pirataria física de CDs sobre o consumo de músicas dos indivíduos. Essa omissão pode gerar um viés sobre os efeitos estimados devido a uma possível relação entre *downloads* e consumo de CDs piratas. Além disso, a literatura empírica existente tipicamente não investiga o efeito potencialmente positivo que o acesso a *downloads* tem sobre o consumo de shows e outros produtos ligados a imagem dos artistas.

Gráfico 1 - Evolução Anual do Índice de Receitas(US\$) da Indústria Fonográfica

Este trabalho pretende preencher estas lacunas ao investigar os efeitos dos *downloads* sobre o consumo de música utilizando uma amostra de estudantes da Universidade de São Paulo. Este estudo apresenta três contribuições à literatura. Em primeiro lugar, é um exercício empírico que, ao nosso conhecimento, é inédito num país em desenvolvimento. Em segundo lugar adiciona-se à análise a influência do consumo de CDs piratas sobre a atividade de *download* e o consumo de música. Finalmente, provemos resultados acerca dos efeitos dos *downloads* sobre o consumo de shows de música que são parte substancial da receita de artistas.

⁵ Segundo uma estimativa da IFPI cerca de 40% dos produtos de áudio e vídeo comercializados no Brasil em 2006 eram de origem ilegal. Fonte: IFPI- The Recording Industry 2006 – Piracy Report – Protecting Creativity in Music. Disponível em: <http://www.ifpi.org/content/library/piracy-report2006.pdf>

Para tanto, coletou-se respostas de 7.147 alunos da universidade de São Paulo por meio de um questionário on-line sobre consumo de música, uso de Internet, opiniões acerca do comércio ilegal, pirataria e características sócio-econômicas⁶.

O grande desafio de se estimar o efeito dos *downloads* sobre o consumo de música utilizando dados em nível individual é a presença de características individuais não observáveis. Em particular, o interesse dos indivíduos por música e tolerância a ilegalidade são características difíceis de serem observadas que podem afetar o consumo de bens musicais. Para lidar com essa dificuldade, utilizamos variáveis instrumentais, uma estratégia comum na literatura que relaciona *downloads* e consumo de CDs (e.g. Rob&Waldfoegel(2006) e Zentner(2006)). Nossos principais instrumentos são variáveis relacionadas à habilidade em informática dos indivíduos bem como as variações individuais sobre opiniões acerca do consumo de bens falsificados. De fato, apresentamos um simples modelo teórico que justifica a adoção dos nossos instrumentos e torna explícitas as hipóteses subjacentes ao seu uso.

Os resultados da análise empírica apontam que a disponibilidade de pirataria afeta significativamente o efeito de *downloads* sobre o consumo de CDs. Individualmente, a prática de *downloads* reduz o consumo de CDs originais em cerca de 25%. Este resultado negativo dos *downloads* sobre o consumo de CDs também é consistente com o resultado obtido numa análise prévia que foi realizada com uma amostra menor de 716 alunos coletada em salas de aula e também mostra evidências de impactos negativos⁷. No entanto, nossos resultados sugerem que tal impacto é minimizado pela substituição de pirataria por *downloads*. Mais acesso a *downloads* gera uma queda no consumo de pirataria, e não apenas no consumo de CDs. Segundo as estimativas deste trabalho, caso não houvesse esta substituição de pirataria por *downloads*, o consumo de CDs apresentaria uma resposta ainda maior à prática de *downloads*: uma redução de cerca de 35%. Evidentemente tais resultados sugerem que a pirataria é um ingrediente importante no mercado de CDs. Segundo nossas estimativas, a prática de pirataria reduz em cerca de 45% o consumo de CDs. Por outro lado, observamos uma compensação (ao menos parcial) deste efeito negativo por efeitos positivos sobre

⁶ Também utilizamos uma amostra secundária, que foi realizada previamente. A amostra consta de 716 observações de alunos de primeiro e segundo ano de 9 institutos da Universidade de São Paulo e foi coletada por meio da aplicação de questionários em salas de aula. Os resultados obtidos com esta amostra são utilizados como base de comparação aos resultados obtidos na amostra principal.

⁷ Estes resultados são fornecidos no apêndice B e sugerem que a prática de *downloads* pela internet reduza em até 37% a probabilidade de comprar CDs. Quando se olha para os impactos sobre a quantidade de CDs consumidos, os resultados mostram que os indivíduos que baixaram CDs, em média compram 3,5 CDs a menos do que os indivíduos que não fazem download de músicas.

consumo de shows, o que levanta a possibilidade de que a Internet também possa ter efeitos positivos sobre a produção de música.

O trabalho é organizado da seguinte maneira: Seção 2 faz um breve resumo da literatura. A seção 3 apresenta um modelo teórico simples que fundamenta a estratégia empírica. A seção 4 faz um detalhamento da amostra e suas estatísticas descritivas. A seção 5 discute os instrumentos utilizados na análise empírica. A seção 6 fornece os resultados gerados bem como as implicações econômicas dos coeficientes estimados. A seção 7 realiza testes de robustez relativos aos resultados obtidos na seção 6. Por fim, a seção 8 discute as conclusões. Um apêndice apresenta tabelas adicionais relativas às análises da seção 6 e 7 e a descrição de uma amostra secundária e os resultados produzidos com esta amostra.

2. LITERATURA

Embora haja um entendimento por parte da Indústria fonográfica de que os *downloads* de música influenciem negativamente a venda de CDs, em termos teóricos o efeito da distribuição de conteúdo protegido sobre o lucro das firmas é ambíguo (Takeyama (1997) e Bakos et ali (1998)). Por um lado, consumidores que baixam músicas pela Internet podem substituir o CD original pela cópia digital. Por outro, os *downloads* de música podem ampliar a divulgação de um artista com um menor custo ao expor os consumidores a novas músicas e artistas que, na ausência da Internet, não seriam revelados.

A ausência de um consenso na literatura teórica motivou o uso de estratégias empíricas com o propósito de um melhor entendimento sobre os efeitos dos *downloads* sobre o consumo de bens associados à música. Contudo, a literatura empírica existente ainda é incipiente. Existem resultados que apontam para existência de um efeito substitutivo dos *downloads* enquanto alguns trabalhos mostram a ausência de qualquer relação entre os *downloads* e o consumo de CDs.

Uma dificuldade central nos trabalhos que utilizam dados em nível individual é que tanto o consumo CDs quanto o de *downloads* são influenciados pelo gosto por música. Gentzkow (2007) aponta como desafio empírico central, quando se procurar estimar a relação de consumo entre dois bens, a tarefa de separar os verdadeiros efeitos substituição ou de complementaridade dos bens da correlação com as preferências dos consumidores. Ou seja, isso reflete o fato de algumas características individuais não observáveis serem correlacionadas com a variável explicada e com as explicativas.

Rob e Waldfogel (2006) utilizam uma amostra de 500 estudantes americanos coletada em 2003 e examinam o efeito dos *downloads* sobre o consumo de CDs e excedente dos consumidores. Utilizando o método de variáveis instrumentais, tendo como fonte de variação exógena para os *downloads* o acesso ou não a conexões de banda-larga das universidades, os autores estimam que cada disco baixado pela Internet reduz em aproximadamente 0,2 unidades de discos comprados em lojas pelos consumidores. Os autores ainda realizam uma análise de valoração de discos e concluem que, geralmente, os álbuns baixados pela Internet

são menos valorizados que os comprados em lojas. Também se conclui que os *downloads* reduzem o gasto de CDs per capita em US\$ 25. Na outra ponta, aumentam o excedente dos consumidores em aproximadamente US\$ 70.

Zentner (2006) utiliza uma base de dados de 15 mil observações em nível individual de 7 países⁸ europeus em 2001 para estimar o efeito de baixar músicas pela Internet sobre a probabilidade de comprar um CD de música. Utilizando medidas de sofisticação do uso de Internet⁹ como instrumentos para resolver o problema de simultaneidade entre *downloads* e consumo de CDs, o autor chega ao resultado de que fazer *download* de músicas implica em uma redução de até 30% na probabilidade de comprar CDs de música. O autor ainda faz uma previsão: se não houvesse a possibilidade de ser fazer *download*, em 2001 os resultados da indústria fonográfica teriam sido 7,8% maiores.

Já Oberholzer-Gee e Strumpf (2007) utilizam dados semanais de *downloads* de música obtidos de um servidor de Internet¹⁰ específico e em seguida, comparam a amostra de *downloads* feitos nos EUA com dados de vendas de CDs (Nielsen-Soundscan) para um número de discos de distintos artistas. O problema de heterogeneidade surge do fato que cada álbum tem um fator específico que influencia sua popularidade. A consequência é que em geral os álbuns mais baixados pela Internet são também os álbuns mais vendidos. Para contornar esse problema de identificação, Oberholzer-Gee e Strumpf (2007) utilizam o método de estimação em dois estágios, utilizando como principal instrumento o número de estudantes alemães que se encontram em período de férias numa determinada semana. A justificativa dos autores é que os arquivos de músicas oriundos da Alemanha representam um terço da oferta dos arquivos disponíveis para *download* aos norte-americanos e quanto maior for o número de estudantes de férias maior será a oferta desses arquivos. Basicamente, os autores partem do pressuposto de que ao entrarem de férias os estudantes alocam mais tempo on-line em programas de transferência de *download*. Sendo assim, o instrumento representa um choque de oferta exógeno de arquivos de música, que implica numa redução do custo de

⁸ Os dados são observações de indivíduos da França, Alemanha, Itália, Holanda, Espanha, Suécia e Inglaterra.

⁹ Os instrumentos utilizados são: se a pessoa possui uma página de internet pessoal (blog), se participou de leilões on-line, pediu suporte técnico on-line, leu revistas especializadas em informática, há quanto tempo usa e-mail e internet.

¹⁰ Os dados referentes a *downloads* de música utilizados pelos autores são provenientes da rede OpenNap que é uma das redes responsáveis pelas transferências de dados do programa Napster.

oportunidade de fazer um *download* nos EUA¹¹. Os resultados obtidos pelos autores apontam que o efeito dos *downloads* sobre as vendas são insignificantes mesmo após a adição de vários controles, e este resultado é robusto a diferentes especificações.

Blackburn(2004) também utiliza dados semanais de vendas de CDs e de atividade de *downloads* para explicar a queda na venda CDs observada após a introdução do Napster. Para controlar a popularidade dos artistas da amostra o autor utiliza choques exógenos no risco de fazer *download*. Ou seja, o autor utiliza os anúncios de processos contra usuários de redes p2p feitos pela RIAA – *Recording Industry Association of America* – como instrumentos para atividade de *download* e mostram que estes anúncios tiveram um impacto redutivo na atividade de *download*¹². Utilizando essa fonte de variação exógena o autor chega a um resultado em que os *downloads* de música impactam de forma negativa. Entretanto, quando se interage um índice popularidade com a variável de *download* o que se observa é um efeito negativo para os artistas famosos e um efeito positivo sobre artistas emergentes. Os resultados sugerem que o efeito negativo encontrado é devido, em grande parte, aos artistas famosos que na amostra analisada são maioria.

Olhando para os efeitos dos *downloads* sobre shows de música, Mortimer e Sorensen (2005) mostram num modelo de bens complementares que a distribuição de músicas pela Internet pode ter dois efeitos opostos sobre a quantidade de shows de música ofertados pelos artistas. À medida que o preço da música de um artista tende a zero devido à possibilidade de *download* da música pela Internet, um maior número de consumidores se torna familiar com o trabalho do artista o que influencia a demanda por shows e conseqüentemente o preço pelo ingresso do show. Na outra ponta, a possibilidade de baixar músicas pela Internet diminui os incentivos dos artistas a fazerem shows, pois os shows se tornam um meio menos efetivo de aumentar as receitas de um CD recém lançado caso uma fração grande das pessoas baixem as

¹¹Quanto mais arquivos MP3 estão disponíveis numa rede P2P maior se torna a probabilidade de um usuário encontrar determinado arquivo disponível para *download*. Isso por sua vez tem uma influência direta sobre o custo de busca esperado para os usuários.

¹² Bhattacharjee et al (2006a) também confirmam esse resultado. Os autores mostram que anúncios de processos contra usuários tem um efeito redutivo de até 90% sobre usuários assíduos e de até 30% sobre usuários moderados. Não obstante, os autores registram o crescimento substancial na atividade de *download* para uma parcel significativa de usuários de programas p2p. Em parte, esse aumento se deve a atividade de usuários de regiões não afetadas pelos processos.

músicas pela Internet¹³. Para testar qual dos efeitos prevalece os autores utilizam dados semanais de vendas de CDs e shows de 2135 artistas durante um período de 10 anos. Os autores analisam as mudanças na oferta de CDs e Shows que coincidem com o surgimento do Napster em 1999 e verificam se essas mudanças são mais intensas nos cortes de artistas e mercados onde se espera que as atividades de *download* sejam mais prevalentes. Os resultados da análise sugerem que o impacto das apresentações sobre a receita de CDs diminui após o surgimento do Napster e que esse declínio é maior nos mercados onde a atividade de *download* é mais prevalente. Entretanto, as receitas de shows e o número de bandas que entram em turnê crescem após o início das atividades de *download* na Internet assim como espaço de tempo entre um lançamento e outro de artistas já estabelecidos.

Uma explicação alternativa para o crescimento do número de bandas após o advento dos *downloads* de música é provida por Oberholzer-Gee e Strumpf (2009). Os autores argumentam que a indústria musical é um torneio onde, sobre uma gama de artistas, somente uma pequena fração se torna popular e provê quase que a totalidade da renda da indústria. Ou seja, a vida de uma artista é uma loteria onde existe uma pequena probabilidade de ele se tornar famoso. O argumento dos autores é de que, embora os *downloads* de músicas pela Internet possam afetar a receita de artistas famosos, a probabilidade de se tornar famoso não é afetada pelos *downloads* de música. É argumentado inclusive que a Internet aumenta essa probabilidade ao oferecer um mecanismo de divulgação menos custoso do que o tradicional.

¹³ É válido argumentar que essa hipótese só seja válida para artistas com uma participação na receita de CDs alta. Grande maioria dos artistas têm uma participação pequena sobre essas receitas e, em alguns contratos, só recebem essa participação após as vendas superarem o ponto de *break-even*.

3.FUNDAMENTAÇÃO TEÓRICA

Para entender as relações entre *download*, CDs e CDs piratas, utilizamos um modelo simples quadrático sobre as preferências de um consumidor. Por meio de uma função utilidade nós podemos explorar a relação das escolhas dos indivíduos com fatores não observados e derivar a fórmula funcional apropriada para a análise empírica.

Suponha que a preferência do i ésimo consumidor é denotada pela seguinte função utilidade ¹⁴:

$$U(x_i, y_i, w_i, m_i) = g_i(Ax_i + By_i + Cw_i) + \rho_i(Dx_i + Ew_i) + y_i\varepsilon_i + x_i\xi_i + w_i\epsilon_i - \frac{b}{2}w_i^2 - \frac{a}{2}x_i^2 - \frac{1}{2}y_i^2 + c_1x_iy_i + c_2x_iw_i + c_3w_iy_i - \gamma_ix_i + m_i$$

E satisfaça a seguinte restrição orçamentária:

$$py_i + qw_i + m_i \leq R_i$$

Onde:

x_i : quantidade de *download* feito pelo consumidor i tal que $x_i \geq 0$.

y_i ¹⁵: quantidade de CDs originais comprados pelo consumidor i tal que $y_i \geq 0$.

w_i : quantidade de CDs piratas comprados pelo consumidor i tal que $w_i \geq 0$.

m_i : montante monetário relativo ao consumo de outros bens relevantes para o consumidor i tal que $m_i \geq 0$.

g_i : termo de gosto não observável por música dos consumidores que afeta *downloads* e CDs originais e CDs piratas.

ρ_i : termo de tolerância a ilegalidade que afeta o consumo de CDs piratas e *Downloads*.

¹⁴ A forma quadrática para a função utilidade é utilizada, pois é uma forma simples que permite trabalhar com a relação de complementaridade dos bens e ao mesmo tempo analisar a influencia do gosto por música sobre as quantidades demandadas de CDs e *downloads* de maneira linear. O leitor mais atento poderia argumentar que essa função utilidade não satisfaz algumas propriedades básicas como a monotonicidade. Entretanto, poderíamos pensar na função apresentada como uma aproximação de Taylor de segunda ordem de uma função que preserve as propriedades usuais de uma função utilidade bem comportada: monotonicidade, convexidade e transitividade das preferências. Como o objetivo é trabalhar com especificações lineares não há perda de generalidade com o uso da função utilidade quadrática.

¹⁵ As conclusões provenientes da análise deste modelo podem ser facilmente generalizadas para a escolha de demais bens musicais, como shows e bens subordinados, pela simples interpretação de y_i sendo a representação destes bens.

A, B, C, D, E, a, b : são constantes reais.

$\varepsilon_i, \xi_i, \epsilon_i$: são termos idiossincráticos no consumo de CDs, *downloads* e CDs piratas respectivamente.

c_1, c_2, c_3 : são constantes de interação entre os bens da cesta de consumo.

γ_i : é um custo subjetivo¹⁶ de *download*.

p, q : são os preços unitários do CD original e pirata respectivamente.

R_i : é a renda monetária mensal do indivíduo i .

Especial atenção deve ser dada as constantes c_1, c_2 e c_3 , pois caso pelo menos uma delas for negativa, implicará a existência de uma relação de substituição entre dois dos bens da cesta de consumo. Já o termo ρ_i representa uma tolerância à ilegalidade que gera um acréscimo sobre a utilidade marginal do consumo de CDs piratas e *downloads*. Especificamente, indivíduos com certos valores morais não se sentem incomodados ou mesmo culpados em consumir produtos de origem ilegal.

O consumidor i resolve o seguinte problema:

$$\begin{aligned} \max U(x_i, y_i, w_i, m_i) \\ \text{s. a } py_i + qw_i + m_i \leq R_i \end{aligned}$$

Ao Substituir a restrição orçamentária $R_i - py_i - qw_i = m_i$ na função utilidade obtemos as seguintes condições de primeira ordem:

$$y_i: y_i = Bg_i + c_1x_i + c_3w_i - p + \varepsilon_i \quad (1)$$

$$x_i: x_i = \frac{1}{a} [Ag_i + D\rho_i + c_1y_i + c_2w_i - \gamma_i + \xi_i] \quad (2)$$

$$w_i: w_i = \frac{1}{b} [Cg_i + E\rho_i + c_2x_i + c_3y_i - q + \epsilon_i] \quad (3)$$

A equação (1) é exatamente a equação estrutural que apresenta os coeficientes de interesse em nossa análise. A constante c_1 denota o efeito marginal dos *downloads* sobre o consumo de CDs e c_3 o efeito marginal do consumo de CDs piratas. Entretanto, tanto em (1) quanto em (2) e (3), o termo g_i , que é a heterogeneidade não observada do gosto por música, afeta ambas as quantidades de *downloads*, de CDs e CDs piratas. Reescrevendo a equação (1) colocando as

¹⁶ O custo de *download* é dito subjetivo, pois não é um custo monetário propriamente dito. O custo está ligado às barreiras de acesso dos usuários aos *downloads* e ao grau de habilidade e conhecimento da internet que pode facilitar ou dificultar que um *download* seja feito. Nesse sentido ele age mais como um termo de desutilidade no consumo de *download* como um custo que entraria na restrição orçamentária.

escolhas de *download*, compra de CD's originais e piratas explicitamente como função das características individuais chega-se à seguinte especificação:

$$y_i(g_i, \rho_i, \varepsilon_i, \gamma_i, \xi_i, \epsilon_i) = Bg_i + c_1x_i(g_i, \rho_i, \varepsilon_i, \gamma_i, \xi_i, \epsilon_i) + c_3w_i(g_i, \rho_i, \varepsilon_i, \gamma_i, \xi_i, \epsilon_i) - p + \varepsilon_i \quad (4)$$

Note-se que tanto x_i , quanto w_i dependem de g_i . Portanto, caso não se possa controlar pelo gosto musical dos indivíduos, estimativas de c_1 e c_3 por mínimos quadrados ordinários serão viesadas e inconsistentes. Uma versão mais explícita da equação 4 revela que o viés da estimativa dos coeficientes pode ser tanto positivo quanto negativo. Ao resolvermos as equações (2) e (3) com a finalidade de isolar o termo g_i e posteriormente substituímos essa expressão na equação (1) obteremos a seguinte equação:

$$y_i = \hat{\alpha} + \hat{c}_1x_i + \hat{c}_3w_i + \delta_i$$

Onde:

$$\hat{\alpha} = \frac{E\gamma_i + Dq - p(DC - EA)}{(DC - EA) - B(Ec_1 - c_3D)}$$

$$\hat{c}_1 = \frac{c_1(DC - EA) - B(Ea + c_2D)}{(DC - EA) - B(Ec_1 - c_3D)}$$

$$\hat{c}_3 = \frac{c_3(DC - EA) - B(Db + c_2E)}{(DC - EA) - B(Ec_1 - c_3D)}$$

$$\delta_i = \frac{\varepsilon_i(DC - EA) + E\xi_i - D\epsilon_i}{(DC - EA) - B(Ec_1 - c_3D)}$$

O que se percebe dos coeficientes acima é que todos são funções dos efeitos marginais dos fatores não observados, g_i e ρ_i sobre as escolhas de *downloads*, CDs e CDs piratas. Contudo, se desenvolvermos desigualdades entre os coeficientes acima e os verdadeiros coeficientes c_1

e c_3 obteremos expressões que não fornecem intuições claras sob quais condições os coeficientes serão subestimados ou sobreestimados¹⁷.

Intuitivamente, o fato de pessoas que compram CD's e fazem *downloads* serem pessoas que gostam de música pode gerar uma correlação espuriamente positiva entre consumo de CD's e Downloads. Por outro lado, pessoas com maior tolerância a ilegalidade tendem a consumir maior quantidade de pirataria e *downloads*, o que pode afetar indiretamente o consumo de CD's.

Da equação (4) observa-se que, se supormos que γ_i e ρ_i sejam independentes de g_i e ε_i , eles satisfazem as condições necessárias para instrumentalizar as variáveis de *download* e consumo de pirataria¹⁸. Isso desempenhará um papel fundamental em nossa estratégia de identificação: utilizaremos variáveis relacionadas a conhecimentos em informática e qualidade de equipamentos disponíveis como *proxies* para o custo de *download*, e respostas relativas às opiniões sobre compra de produtos ilegais como *proxies* de tolerância à ilegalidade.

Outro ponto importante que podemos inferir com este modelo é que ele mostra como as estimativas podem ser afetadas quando utilizamos uma estratégia que tenta identificar somente o coeficiente de *download* e ignora o consumo de CDs piratas na equação (1). Supondo que a variável de *download* seja instrumentalizada com as variações do custo subjetivo de *download* e, portanto, seja uma função apenas das variações do custo de *download*, por meio da regra da cadeia, podemos enxergar o viés gerado sobre os coeficientes estimados devido à omissão do consumo de CDs piratas. Derivando y_i na equação (1), em relação à γ_i , temos:

$$\frac{\partial y_i}{\partial \gamma_i} = \frac{\partial y_i}{\partial x_i} \frac{\partial x_i}{\partial \gamma_i} + \frac{\partial y_i}{\partial w_i} \frac{\partial w_i}{\partial x_i} \frac{\partial x_i}{\partial \gamma_i} = -\left(\frac{c_1}{a} + \frac{c_2 c_3}{ab}\right)$$

¹⁷ Basicamente, se supormos que: $\widehat{c}_1 > c_1 \rightarrow c_1^2 - \frac{D}{E} c_1 c_3 + a < \frac{D}{E} c_2$ e analogamente, $\widehat{c}_3 > c_3 \rightarrow c_3^2 - \frac{E}{D} c_1 c_3 + b > -\frac{E}{D} c_2$.

¹⁸ Trivialmente pode-se mostrar que ε_i e g_i aparecem separadamente a ρ_i e γ_i na determinação de $x_i(\cdot)$ e $w_i(\cdot)$, o que qualifica ρ_i e γ_i como instrumentos.

Como o coeficiente dos *downloads* é representado pelo efeito marginal dos *downloads* na equação (1) temos que:

$$\frac{\partial y_i}{\partial x_i} = \frac{\frac{\partial y_i}{\partial \gamma_i}}{\frac{\partial x_i}{\partial \gamma_i}} = \frac{-\left(\frac{c_1}{a} + \frac{c_2 c_3}{ab}\right)}{-\frac{1}{a}} = c_1 + \frac{c_2 c_3}{b}$$

A expressão acima mostra que se ignorarmos o consumo de CDs piratas e os coeficientes c_2 e c_3 possuírem o mesmo sinal, então, o coeficiente de *downloads*, caso seja negativo, será subestimado em termos absolutos. Quando se omite o consumo de CDs piratas na equação (1), as variações no custo subjetivo de *download* não são exógenas devido à relação dos *downloads* com a pirataria física de CDs que, por ser omitida, se torna uma componente do erro aleatório¹⁹. Este é um comportamento previsto pelo modelo que é observado empiricamente na seção de resultados.

¹⁹ Intuitivamente, se supormos uma relação de substituição entre pirataria, download e CDs, quando não controlamos o consumo de pirataria, a ausência do consumo de CDs de um indivíduo que consome CDs piratas e não faz *download* contribui para uma amenização da relação de causalidade entre downloads e CDs.

4. DADOS E ESTATÍSTICAS DESCRITIVAS

Este trabalho é baseado em uma amostra de dados coletados por meio de uma pesquisa on-line. Inicialmente, construiu-se uma página na Internet onde o questionário poderia ser respondido. Em seguida foi enviado por email um convite a todos os alunos de graduação e pós-graduação da Universidade de São Paulo – USP, solicitando o preenchimento do questionário. Durante o período de 6 dias, 7.147 pessoas responderam o questionário gerando uma amostra de aproximadamente 8% da população de alunos matriculados na USP²⁰.

Um possível problema com este tipo de amostragem é que existe uma seleção no sentido de que algumas pessoas deixam de participar da pesquisa devido a fatores não observáveis. Um exemplo claro é o fato de um indivíduo selecionado não ter paciência em preencher formulários ou mesmo não ter tido tempo para isso. Sendo assim, esta amostra seleciona pessoas que possuem uma maior propensão a participar de pesquisas. Não obstante, espera-se que esta propensão não esteja correlacionada com o gosto por música dos participantes que é a principal fonte de endogeneidade de nossa análise. Uma evidência nesse sentido emerge quando se compara a distribuição da preferência por estilos musicais com a distribuição da mesma variável obtida numa amostra preliminar coletada em salas de aulas²¹ da Universidade de São Paulo. O gráfico abaixo mostra que as preferências entre os estilos musicais não são afetadas pela forma como a seleção da amostra foi feita. Um teste de Komogorolv - Smirnov²² é feito para testar se as distribuições de estilos musicais são estatisticamente diferentes. O valor-p da estatística do teste é de 0,87 e sugere que a hipótese nula, de que as distribuições não são diferentes, não seja rejeitada.

²⁰ Segundo o anuário estatístico da USP em 2009 a universidade possuía 88.261 alunos matriculados. Nesse número inclui-se alunos de graduação e pós-graduação (mestrado e doutorado). (Fonte: USP em Números. Disponível em: http://sistemas3.usp.br/anuario/usp_em_numeros.pdf.

²¹ No apêndice A.2 é fornecido algumas das estatísticas descritivas a respeito desta amostra.

²² O teste de Komogorov-Smirnov é utilizado para determinar se duas distribuições de probabilidade diferem uma da outra ou se uma das distribuições de probabilidade subjacentes difere da distribuição em hipótese, em qualquer dos casos com base em amostras finitas.

Gráfico 2 - Distribuição dos Estilos Musicais

Outro ponto importante da amostra em questão (feita por Internet) é a existência de variabilidade entre os respondentes da amostra. Observamos respostas de indivíduos de todos os institutos e faculdades integrantes da Universidade de São Paulo. Isto é uma evidência de que os fatores que influem na seleção não são características exclusivas de um grupo específico da população em estudo.

Ademais, o questionário utilizado nesta análise inclui informações sobre a atividade de *download*, consumo de CDs, CDs piratas, Shows de música e Bens Subordinados (produtos ligados a imagem do artista ou banda, mas que não são considerados como um produto musical²³) referentes ao período de 6 meses anteriores a data de aplicação da pesquisa. Também captamos informações específicas sobre o conhecimento de informática dos indivíduos, consumo de produtos comercializados em feiras e camelôs e informações sobre a opinião sobre o comércio de tais produtos. Outras variáveis captadas são relativas a informações demográficas como idade, sexo, renda, raça, anos de estudo na universidade e instituto de matrícula na USP. No apêndice D fornecemos uma cópia do questionário utilizado para a pesquisa.

A idade média dos participantes é de quase 25 anos, cerca de 96% possuem acesso à Internet em casa, 75% possuem um laptop e 85% afirmam fazer *download* de músicas pela Internet.

²³ Alguns exemplos de Produtos Subordinados são: Camisetas de Banda, Adesivos, Bandeiras com o nome da Banda ou Artista. Embora não sejam considerados como um produto musical estes produtos financiam a produção artística devido ao pagamento de *royalties* do uso da imagem dos artistas.

Este último dado mostra que a atividade de *download* é algo comum entre estudantes e também é consistente com os dados de uma pesquisa de mercado²⁴ da ABPD- *Associação Brasileira dos Produtores de Discos*. Esta pesquisa mostra que das pessoas que afirmam ter feito *download* em dez regiões metropolitanas brasileiras, 69% possuem pelo menos o segundo grau completo. Sendo assim, como a atividade de *download* é realizada, em grande parte, por pessoas com maior grau de instrução, utilizar uma amostra de estudantes universitários se mostra apropriado para estudar os efeitos dos *downloads* sobre o consumo de música. A Tabela 1, abaixo, fornece as principais estatísticas sumárias da amostra.

Tabela 1 - Estatísticas Sumárias

<i>Estatísticas</i>	
<i>Amostra Total</i>	N=7,147
<i>idade média</i>	24.7
	%
Pessoas do Sexo Masculino	46.9
Pessoas que baixam música pela internet	85.12
Pessoas com conexão à internet em casa	96.6
Possuem Laptop com conexão wireless	75.5
Pessoas que compraram CD nos últimos 6 meses	34.0
Pessoas que foram a Shows nos últimos 6 meses	62.7
Pessoas que compraram Bens subordinados	17.58
Pessoas com conexão à internet (banda-larga)	93.4
Pessoas que compram CDs piratas	16.5
<i>Pessoas que fazem download de músicas</i>	N=6,069
	%
Pessoas que compraram CDs nos últimos 6 meses	32.2
Pessoas com conexão à internet (banda-larga)	94.8
Pessoas que consomem CD pirata	16.36
Pessoas que foram a Shows nos últimos 6 meses	65.0
<i>Pessoas que não fazem download de músicas</i>	N=1,061
	%
Pessoas que compraram CDs nos últimos 6 meses	46.5
Pessoas com conexão à internet (banda-larga)	85.7
Pessoas que consomem CD pirata	17.15
Pessoas que foram a Shows nos últimos 6 meses	50.4

Quando olhamos para as pessoas que fazem *download* de música observamos que a proporção delas que comprou um CD é menor do que a proporção observada na amostra restrita às pessoas que não fazem *download*. A despeito do gosto por música não observado, essa comparação já mostra uma evidência de que existe um comportamento substitutivo de

²⁴ A pesquisa foi realizada em 2006 e conta com uma amostra de 1.209 pessoas de 10 regiões metropolitanas Brasileiras: São Paulo, Rio de Janeiro, Ribeirão Preto, Belo Horizonte, Porto Alegre, Curitiba, Salvador, Fortaleza, Recife e Brasília. Disponível em: http://www.abpd.org.br/musicaInternet_pesquisa.asp

downloads e CDs. Por outro lado, com a mesma base de comparação vemos que a proporção de pessoas que foram a shows de música é 15 pontos percentuais maior entre as pessoas que utilizam a Internet para baixar músicas. Já o consumo de CDs piratas não varia muito entre os estratos da amostra permanecendo em cerca de 16% a 17 % entre os estratos. A tabela abaixo reforça essas evidências fornecendo resultados de testes de Komogorolv-Smirnov entre as subamostras de indivíduos que fazem *download* e de indivíduos que não fazem *download*.

Tabela 2 - Testes de Komogorolv - Smirnov

H0: As distribuições entre as pessoas que baixam músicas não são diferentes	
	Valor - P*
Consumo de CDs	0.000
Nº de Shows	0.000
Consumo de CDs piratas	0.311
Consumo de Bens Subordinados	0.002

*Valor P da Estatística do Teste Komolgorov - Smirnov

Os resultados da tabela acima confirmam que existe uma diferença estatisticamente significativa entre as distribuições de Consumo de CDs, Shows e Bens Subordinados quando comparamos as subamostras de pessoas que fazem *download* com a das pessoas que não fazem *download* de música. O teste também confirma que as distribuições do consumo de CDs piratas não se distinguem significativamente.

Não obstante, essa primeira análise não leva em conta a influência que o gosto por música tem sobre as escolhas dos indivíduos. O gráfico abaixo mostra a relação entre os níveis de interesse declarado²⁵ por música com a proporção de indivíduos da amostra que já consumiram CDs, CDs piratas, shows, bens subordinados e fizeram *downloads* de música. Embora o interesse declarado não seja uma medida precisa do gosto por música, no gráfico observamos que existe uma relação crescente da propensão de consumo desses bens condicional aos níveis de interesse declarado. A exceção fica por conta da proporção de pessoas que comprou CDs piratas que parece não ser afetada positivamente pelo interesse em música. Entretanto, como salientado pelo modelo teórico, o gráfico reforça o impacto

²⁵ No questionário solicitamos aos respondentes que assinalassem dentre um conjunto de quatro opções aquela que melhor expressasse o interesse por música do respondente. Dentre as opções os respondentes poderiam marcar: “*não interessado em música*”, “*pouco interessado em música*”, “*interessado em música*” e “*muito interessado em música*”.

simultâneo dessa característica individual sobre as escolhas dos indivíduos e deixa explícito que numa análise mais rigorosa devemos de alguma forma controlar essa simultaneidade.

Gráfico 3 - Relação do Interesse Declarado por música com a propensão de consumo de Música

Sabendo deste problema, a análise que se segue apóia-se no uso de variações de *downloads* devidas ao uso de variáveis instrumentais. Um ponto importante a se ressaltar, é que das pessoas que não baixam música pela Internet, cerca de 86% possuem conexão à Internet do tipo banda-larga. Isso por sua vez fornece uma evidência de que o comportamento de *download* não é influenciado apenas pelo simples acesso à Internet. Outros fatores como a habilidade em informática e mesmo a tolerância a ilegalidade podem ter algum tipo de influência sobre o custo de oportunidade ou sobre a utilidade percebida na atividade de *download* de músicas e consumo de CDs piratas.

5. ESTRATÉGIA DE IDENTIFICAÇÃO

Como vimos acima e no modelo desenvolvido o gosto por música influencia simultaneamente as escolhas de CDs, CDs piratas e *downloads*. Numa versão modificada do modelo pode-se mostrar que esse gosto por música também influencia as escolhas de shows de música. Para identificar corretamente os efeitos dos *downloads* e consumo de CDs piratas sobre as escolhas de consumo de CDs e Shows, utilizamos variáveis instrumentais que se associam ao custo subjetivo de *download* e a tolerância a ilegalidade dos indivíduos.

Um primeiro conjunto de instrumentos que é utilizado neste trabalho são variáveis ligadas a habilidade de informática dos indivíduos²⁶. Esperamos que as pessoas que possuam um maior conhecimento sobre informática tenham um menor custo subjetivo para baixar músicas pela Internet e com isso tenham uma maior probabilidade de baixar músicas que os indivíduos leigos em informática.

Para mensurar tal habilidade no questionário, duas perguntas solicitam que os respondentes assinalem opções acerca dos itens da configuração do computador pessoal que o indivíduo conhece e atividades possíveis de serem realizadas no computador. Os itens da configuração basicamente se referem à memória RAM- *Random Access Memory*, memória em disco rígido (HD- *Hard Drive*) e o processador utilizado no computador. Já as atividades de informática perguntam se o indivíduo já desfragmentou o disco, utilizou antivírus, formatou o computador, instalou *software* e se já instalou algum componente de *hardware* em seu computador pessoal. Com essas perguntas são geradas variáveis *dummies* que indicam se o indivíduo possui ou não as características que cada pergunta busca captar. Essas *dummies* são utilizadas na equação de primeiro estágio como variáveis instrumentais e mensuram o efeito marginal que cada item tem sobre a probabilidade de baixar músicas²⁷.

²⁶ Numa análise prévia, realizada com a amostra feita em salas de aula, utilizamos variáveis ligadas que buscavam mensurar a habilidade de navegação na internet do indivíduo. Basicamente, duas variáveis eram utilizadas com instrumentos. A taxa de conclusão de downloads de arquivos e o Número de programas e sites já utilizados pelo indivíduo para fazer download. O problema com essas variáveis é que as escolhas dos níveis dessas variáveis possivelmente podem ser influenciadas pelo gosto por música dos indivíduos.

²⁷ Outro tipo de especificação possível é a agregação das respostas para cada indivíduo. Nesse caso seriam geradas duas variáveis. Uma mede o número de itens da configuração do computador que o indivíduo conhece e outra o número de atividades que ele já desempenhou num computador. Esperar-se-ia que quanto maior fosse o número observado, maior seria a habilidade em informática do indivíduo. Não obstante, a especificação mais correta a ser utilizada é a das variáveis *dummies* para cada tipo de resposta. A agregação das respostas pode gerar um erro de medida devido a não comparabilidade da composição da resposta de cada indivíduo. Como

Outro conjunto de instrumentos está associado às variações na tolerância a ilegalidade, que podem explicar o consumo de CDs piratas (e possivelmente *downloads*) sem afetar diretamente as escolhas de CDs originais. Apesar de ser uma característica não observada, pelo questionário conseguimos captar algumas opiniões acerca do consumo de bens falsificados que podem refletir essa tolerância. Especificamente, em uma das perguntas do questionário, é perguntado se os indivíduos são contra ou favor do comércio de produtos de origem clandestina ou ilegal. Pessoas que se auto declaram contra a pirataria possivelmente são pessoas com uma menor tolerância a ilegalidade.

Outra variável que pode captar essa tolerância é uma variável que, dentre uma lista de produtos ofertados em feiras e mercados informais, soma aqueles produtos que o indivíduo já adquiriu por meio desses mercados²⁸. Entre estes produtos perguntamos se o indivíduo já comprou acessórios de aparelhos eletrônicos, bolsas, celular, baralho, jóias, mochila, produtos esportivos, relógios, roupas, sapato, tênis e óculos. Quanto maior a soma de produtos, presume-se que maior seja a exposição do indivíduo a mercados informais e, portanto, maior seja a probabilidade de consumo de produtos ilegais o que revela uma baixa tolerância à ilegalidade.

A tabela abaixo fornece algumas estatísticas descritivas acerca dos instrumentos utilizados em nossa análise:

cada tipo de resposta pode ter uma influência específica sobre a probabilidade de fazer download, então um indicador que revela que o indivíduo fez duas atividades ligadas à informática pode não ser comparável com um mesmo indicador de outro indivíduo que também fez duas atividades de informática.

²⁸ Importante ressaltar que na soma desses produtos são excluídos os produtos cuja escolha provavelmente são atribuídas ao gosto musical dos indivíduos. Ex: DVD's, CD's, aparelhos de áudio e vídeo-games.

Tabela 3 - Estatísticas Descritivas dos Instrumentos

Instrumentos	N	Média
Configuração	7147	2,25
Memória hd	7147	0,77
Memória ram	7147	0,73
Processador	7147	0,74
Atividades de informática	7147	3,33
desfragmentou disco	7147	0,68
formatou computador	7147	0,55
instalou software	7147	0,82
instalou hardware	7147	0,35
rodou antivírus	7147	0,91
Nº de Produtos de Feira ou Camelô	7147	1,08
É a favor da pirataria	7097	0,16
É Contra a pirataria	7097	0,43

Nota:

hd = conhece a memória HD do computador
desfragmentador= já desfragmentou o disco do computador
formatador = já formatou o computador
instalar software = já instalou software em computador
antivirus = já utilizou antivirus no computador

Outro ponto relevante acerca dos instrumentos é sua não correlação com o fator de heterogeneidade não observada que é o gosto por música. Apesar de não observamos perfeitamente o gosto por música, podemos testar essa possível correlação regredindo os instrumentos utilizados com variáveis que possam refletir um gosto por música mais acentuado. Uma maneira de checar se os instrumentos são correlacionados com o gosto por música é regredi-los com relação ao interesse por música declarado dos indivíduos e uma *dummy* que revela se o indivíduo já tocou ou toca algum instrumento musical. A tabela abaixo mostra tal exercício empírico.

O propósito deste exercício é verificar se o gosto por música tem alguma influência sobre a probabilidade de indivíduo ter executado alguma das atividades de informática. Poderia se pensar na regressão inversa, entretanto, é provável que os instrumentos se correlacionem com o gosto por música devido a sua provável relação com *downloads* de música que podem influenciar o gosto por música. A princípio essa é uma relação desconhecida. Ao fazer *downloads* os indivíduos podem começar a desenvolver seu gosto por música ou por outra razão deixar de gostar de música.

As regressões estimadas na tabela abaixo excluem a faixa “ *muito interessado em música*”, ou seja, a interpretação dos coeficientes de interesse declarado por música é em relação ao maior nível de interesse. As cinco primeiras especificações (colunas) analisam a correlação das atividades de informática com o gosto por música enquanto que as últimas três colunas mostram os coeficientes para os itens da configuração do computador.

Um padrão que se observa é que as *dummies* ligadas atividades de informática são influenciadas pela *dummie* de instrumento musical e os itens da configuração pelas *dummies* de interesse declarado por música. Embora significantes, as magnitudes dos coeficientes da *dummie* “*Toca instrumento musical*” são pequenas. A interpretação que obtemos é que tocar ou ter tocado algum instrumento musical aumenta em cerca de 1% a 4% a probabilidade de ter executado alguma atividade de informática. Entretanto, é válido ressaltar que a significância desses coeficientes pode ser devida a algum fator não observado ligado a habilidades cognitivas dos indivíduos. Indivíduos mais habilidosos poderiam tanto ter uma maior facilidade para tocar instrumentos musicais quanto para aprender a utilizar um computador e isso poderia gerar uma relação espúria entre as duas variáveis. Outra possível explicação seria o fato de que os músicos geralmente utilizam computadores em suas rotinas. Seja para gravar uma composição ou mesmo pesquisar sobre partituras ou composições de outros artistas o uso de computadores pode estar ligado a essas atividades e, portanto, ter tocado algum instrumento pode influenciar o conhecimento de informática.

Tabela 4- Relação dos Instrumentos com Interesse Declarado por Música – Mínimos Quadrados Ordinários

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variáveis Explicativas	Desfragmentou o Computador	Formatou Computador	Utiliza Antivírus	Instalou Software	Instalou Hardware	Contra Pirataria	Produtos de Camelô ou Feira	Memória RAM	Memória HD	Processador
Toca Instrumento Musical	0.0330*** (0.0106)	0.0301*** (0.0114)	0.0122* (0.00895)	0.0177** (0.00888)	0.0445*** (0.0103)	0.00278 (0.0123)	0.0568 (0.0405)	0.0206** (0.0103)	0.0105 (0.0101)	0.00387 (0.0104)
Não Interessado em música	0.00274 (0.0458)	0.0302 (0.0495)	-0.0479 (0.0362)	-0.0537 (0.0400)	-0.00784 (0.0525)	-0.0427 (0.0578)	0.0834 (0.203)	-0.128*** (0.0488)	-0.155*** (0.0501)	-0.140*** (0.0505)
Pouco Interessado em música	0.00381 (0.0153)	-0.0198 (0.0164)	0.0109 (0.0102)	-0.00542 (0.0125)	-0.0212 (0.0151)	-0.0158 (0.0180)	0.0320 (0.0608)	-0.0918*** (0.0150)	-0.0906*** (0.0147)	-0.0753*** (0.0152)
Interessado em Música	-0.00386 (0.0120)	-0.0131 (0.0130)	0.000948 (0.00771)	-0.0156 (0.00959)	-0.0448*** (0.0123)	-0.0252* (0.0142)	-0.0147 (0.0484)	-0.0348*** (0.0114)	-0.0337*** (0.0110)	-0.0224* (0.0115)
Observações	7003	7003	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.208	0.195	0.080	0.178	0.274	0.046	0.057	0.159	0.130	0.136

Desvios Padrões Robustos entre parênteses

*** p<0.01, ** p<0.05, * p<0.1

Variáveis controles: idade, sexo, cor, renda, instituto e Ano de Ingresso na Universidade

Já com relação às variáveis ligadas aos itens da configuração do computador percebe-se que há uma relação crescente das *dummies* de interesse declarado por música com a probabilidade de conhecer estes itens. Quando olhamos da faixa “interessado em música” para faixa “não interessado em música” observa-se um aumento do coeficiente estimado o que implica que as pessoas não interessadas em música têm uma probabilidade menor de saber um item da configuração do computador em relação às pessoas muito interessadas. Isso por sua vez pode ser um indício da correlação destes instrumentos com o gosto por música dos indivíduos. Intuitivamente, pode-se pensar que os *downloads* de música dependem também da qualidade do equipamento utilizado. Como indivíduos mais interessados em música também são indivíduos que baixam mais músicas, as características de um computador podem ser relevantes para indivíduos que pretendem adquirir um computador com o intuito de baixar músicas.

Não obstante, é válido lembrar que essas faixas de interesse por música são medidas ordinais do interesse por música. Ou seja, elas são não comparáveis entre os indivíduos e isso pode gerar um erro de medida. Basicamente, muito interessado em música para uma pessoa pode ser apenas pouco interessado para outra. Isso por sua vez, indica que não podemos saber com precisão se existe uma correlação dos instrumentos com as medidas de interesse.

Contudo, mesmo que essa correlação com o interesse declarado por música seja de fato consistente nós observamos essa medida de gosto e podemos controlá-la na equação estrutural. O grande ponto é que estas medidas não são perfeitas e possivelmente outros fatores, que compõe o gosto por música e são não observáveis, influenciam as decisões dos indivíduos com respeito ao consumo de música. A influência destes fatores sobre os instrumentos não pode ser diretamente testada e nesse caso se recorre às evidências fornecidas pelos resultados dos testes de sobreidentificação²⁹ fornecidos na seção de resultados.

²⁹ Sob a hipótese nula de que os instrumentos são válidos a estatística J-Hansen do teste de sobreidentificação $Q(\hat{\beta}) = \left\{ \frac{1}{N} (Y - X'\hat{\beta})'Z \right\} \hat{S}^{-1} \left\{ \frac{1}{N} Z'(Y - X'\hat{\beta}) \right\}$ tem distribuição qui-quadrado com graus de liberdade igual ao número de restrições de sobreidentificação. Como a hipótese de exogeneidade dos instrumentos é de que eles sejam ortogonais ao erro, temos que: $E[Z'(Y - X'\hat{\beta})] = 0$. Portanto, sob a hipótese nula a estatística do teste de sobreidentificação é igual a zero.

Ademais, espera-se que a direção do viés gerado nos coeficientes devido à correlação dos instrumentos com o gosto por música seja de subestimação. Numa relação entre *downloads* e CDs, caso os instrumentos fossem influenciados por fatores ligados ao gosto por música que não podem ser controlados na equação estrutural, deveríamos observar uma minimização dos efeitos de *download*. Entretanto, o que se observa na próxima seção é que quando utilizamos o método de variáveis instrumentais nota-se um aumento dos efeitos dos *downloads* sobre o consumo de CDs.

Por fim, as colunas (6) e (7) mostram que não há um padrão significativo entre o interesse por música declarado e as variáveis ligadas a tolerância a ilegalidade e consumo de bens ofertados em mercados informais.

5.1 Resultados de Primeiro Estágio

A hipótese crucial acerca dos instrumentos utilizados é que eles podem afetar o custo de oportunidade de baixar músicas ou mesmo revelar algum padrão de desutilidade no consumo de bens falsificados que afetem a probabilidade de fazer *downloads* e consumir CDs piratas. Nas estimativas de primeiro estágio consideramos o seguinte modelo empírico:

$$d_i = X_i' b_{11} + Z_i' b_{12} + \varepsilon_{1i} \quad (5)$$

$$w_i = X_i' b_{21} + Z_i' b_{22} + \varepsilon_{2i} \quad (6)$$

Onde d_i e w_i são variáveis *dummies* que indicam se a pessoa utiliza ou não a Internet para baixar músicas e se comprou ou não CDs piratas nos últimos 6 meses respectivamente. O vetor X_i representa os controles que são: idade, sexo, anos de estudo na Universidade, cor ou raça, instituto em que estuda na Universidade, renda familiar em reais e interesse declarado por música. O vetor Z_i representa os instrumentos utilizados nas especificações de primeiro estágio.

Note que são estimadas duas equações distintas de primeiro estágio. A equação (5) procura estimar os efeitos sobre a probabilidade de utilizar a Internet para baixar músicas devido a

variações nos instrumentos utilizados enquanto que a equação (6) mensura os mesmos efeitos sobre a probabilidade de comprar CDs piratas.

Abaixo, a tabela 5 mostra os coeficientes referentes ao vetor Z_i das equações (5) e (6) por mínimos quadrados ordinários. Na tabela utilizamos 3 especificações de primeiro estágio. A primeira delas exclui os instrumentos ligados ao consumo de CDs piratas enquanto que as duas últimas alternam o uso dos instrumentos ligados a tolerância a ilegalidade.

Quando olhamos para as estimativas de primeiro estágio das colunas (1), (2) e (3) vemos que alguns dos itens da configuração não mostram significância estatística sobre a probabilidade de utilizar a Internet para baixar músicas. Somente a *dummie* ligada ao conhecimento da memória do disco rígido do computador é significativa implicando um aumento de 7% sobre a probabilidade de fazer *download*. As *dummies* relacionadas à atividade de informática são significantes e implicam em incrementos de cerca de 2,5% a 4% sobre a probabilidade de fazer *download*. Estes primeiros resultados implicam que um indivíduo que tenha um conhecimento pleno de informática pode ter uma probabilidade até 18% maior de fazer *download* do que um indivíduo que é leigo em informática.

Na outra ponta, nas colunas (4) e (5) vemos que as variáveis ligadas ao conhecimento de informática possuem um impacto negativo, entretanto, não significativa sobre a probabilidade de ter comprado um CD pirata. Essa não significância estatística dos instrumentos ligados a conhecimento de informática sobre o consumo de pirataria é consistente a com especificação do modelo teórico, onde o custo subjetivo de *download* afeta somente as quantidades de *downloads* de música.

Já as variáveis ligadas a tolerância a ilegalidade são significantes tanto para explicar a probabilidade de fazer *download* quanto à de ter comprado um CD pirata. Observa-se que as magnitudes dos coeficientes estimados são maiores quando essas variáveis são utilizadas para explicar o consumo de pirataria. Os coeficientes das colunas (2) e (4) mostram que em média as pessoas que são contra a pirataria têm uma probabilidade cerca de 10% menor de comprar CDs piratas e 4% menor de fazer *downloads* de músicas. Comparando estes resultados com o modelo teórico vemos que o coeficiente E - denota a utilidade marginal de tolerância a ilegalidade devida ao consumo de CDs pirata - é maior do que D - denota a utilidade marginal de tolerância a ilegalidade devido aos *downloads* de música.

Tabela 5 - Estimativas de Primeiro Estágio - Mínimos Quadrados Ordinários

Variáveis Instrumentais	Utiliza a Internet para fazer Download de música			Comprou CD pirata	
	(1)	(2)	(3)	(4)	(5)
Desfragmentou o computador	0.0381*** (0.0106)	0.0361*** (0.0106)	0.0378*** (0.0106)	-0.0165 (0.0113)	-0.0128 (0.0114)
Formatou o computador	0.0269*** (0.00932)	0.0259*** (0.00928)	0.0253*** (0.00931)	0.000678 (0.0103)	-0.000599 (0.0103)
Utiliza Antivirus	0.0418** (0.0177)	0.0413** (0.0176)	0.0408** (0.0176)	-0.0193 (0.0176)	-0.0205 (0.0178)
Conhece a Memória Ram	0.00364 (0.0176)	0.00362 (0.0176)	0.00468 (0.0176)	-0.0297 (0.0185)	-0.0274 (0.0187)
Conhece o Disco Rígido (HD)	0.0714*** (0.0192)	0.0718*** (0.0191)	0.0705*** (0.0191)	0.0186 (0.0194)	0.0160 (0.0195)
Conhece o Processador	-0.000588 (0.0153)	0.000504 (0.0152)	-0.000288 (0.0152)	-0.0107 (0.0171)	-0.0125 (0.0173)
É contra a Pirataria		-0.0474*** (0.00833)		-0.103*** (0.00883)	
Nº de Produtos de Camelô ou Feira		0.00719*** (0.00231)	0.0106*** (0.00228)	0.0319*** (0.00332)	0.0394*** (0.00321)
Constante	1.115*** (0.0372)	1.109*** (0.0373)	1.095*** (0.0374)	0.145*** (0.0367)	0.115*** (0.0369)
Estatística F (instrumentos excluídos) ¹	20.13	21.60	20.46	41.52	22.99
Observações	7003	7003	7003	7003	7003
R ²	0.156	0.162	0.158	0.076	0.059

Desvios Padrões robustos entre Parênteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP

e Interesse declarado por música

¹ Teste F de significância conjunta dos instrumentos utilizados

Para cada especificação de primeiro estágio é fornecido o valor das estatísticas F de significância global dos instrumentos utilizados. Utilizando a regra de bolso de Staiger e Stock (1997)³⁰ nós temos boas evidências de que os instrumentos utilizados não são fracos. Em todas as especificações de primeiro estágio as estatísticas calculadas estão bem acima do valor de corte que é 10.

No apêndice C, a tabela C.1 fornece estimativas para outras três medidas de *download*, que são: Se a pessoa baixou pelo menos um CD nos últimos 6 meses, se a pessoa baixou pelo menos uma música avulsa pela Internet nos últimos 15 dias, se a pessoa faz *download* ao menos uma vez na semana. Mesmo com diferentes medidas de *download* os instrumentos mantêm o padrão apresentado na tabela acima o que mostra uma boa evidência de que estes resultados são consistentes.

³⁰ Segundo os autores uma estatística F menor do que 10 é uma evidência de que os instrumentos utilizados são fracos. O problema de se utilizar instrumentos com pouco poder explicativo sobre as variáveis endógenas é que as estatísticas geradas na equação estrutural já não têm tanto suporte na teoria assintótica. Isso por sua vez, gera um problema grave de inferência caso seja ignorado o problema com as variáveis instrumentais.

6. ABORDAGEM EMPÍRICA

Na seção de estatísticas descritivas uma simples comparação de médias mostra que, dentre as pessoas que fazem *download*, existe uma maior proporção de pessoas que foram a shows de música e uma menor proporção de pessoas que compraram CDs. Por sua vez, isto sugere que talvez os *downloads* tenham um efeito negativo sobre a propensão de consumo de CDs e positivo sobre shows de música. Para testar essas evidências é considerado o seguinte modelo:

$$y_i = c_1 d_i + c_3 w_i + X_i' b + \varepsilon_i \quad (7)$$

$$d_i = X_i' b_{11} + Z_i' b_{12} + \varepsilon_{1i} \quad (5)$$

$$w_i = X_i' b_{21} + Z_i' b_{22} + \varepsilon_{2i} \quad (6)$$

Onde para indivíduo i , y_i , d_i , w_i são variáveis discretas³¹ $\{0,1\}$ indicando que, nos últimos 6 meses, o indivíduo i comprou um CD de música³², baixou músicas e comprou um CD pirata respectivamente. O vetor X_i representa as características individuais³³ observáveis, Z_i representa o vetor de variáveis instrumentais discutidas na seção anterior e ε_i representa o erro.

Para estimar a equação (10) são considerados dois métodos. O primeiro consiste nos mínimos quadrados ordinários (MQO) e o segundo utiliza variáveis instrumentais que leva em conta a estimação das equações (5) e (6) como primeiros estágios. Especificamente, é utilizado o método generalizados dos momentos (GMM) cujas estimativas, em termos assintóticos, são consistentes, robustas à heteroscedasticidade, normalmente distribuídas e, para uma escolha ótima da matriz de ponderação, são as mais eficientes dentre os estimadores lineares de

³¹ A razão para utilizarmos apenas variáveis indicadoras tanto nas variáveis dependentes quanto explicativas é o erro de medida gerado quando se reporta quantidades consumidas de CDs, downloads e CDs piratas. Como essas variáveis representam contagens feitas pelos próprios participantes da pesquisa é provável que essas medidas não sejam medidas exatas dos consumos realizados. Não obstante, o uso de variáveis dicotômicas pode ser uma maneira de eliminar esse erro de medida.

³² A variável y_i também pode ser usada para representar se i foi a um show de música.

³³ São utilizados controles como a renda familiar que é controlada por faixas que vão de um mínimo de “menos que 900 reais” até um máximo de “acima de 15.000 reais”, outros controles como idade, sexo, cor, instituto da universidade e anos de estudo na USP são adicionados a formula funcional da equação (10).

variáveis instrumentais³⁴. A comparação destes dois métodos de estimação permite um melhor entendimento da direção do viés gerado devido à omissão do gosto por música na equação estrutural.

A tabela 6 mostra os coeficientes estimados da equação (7) para os dois métodos considerados. Nas colunas (1) e (2), empregamos o método de mínimos quadrados generalizados para controlar uma possível heteroscedasticidade entre os indivíduos. A coluna 1 mostra um efeito negativo do uso da Internet para baixar músicas. Fazer *download* de músicas implica numa probabilidade de comprar CDs cerca de 10% menor do que os indivíduos que não fazem *download*. Entretanto, esse resultado não leva em conta uma possível influência de fatores heterogêneos não-observados. A coluna (2) da tabela 6 apresenta uma estimativa por mínimos quadrados ordinários que leva em conta as *dummies* de interesse declarado por música dos indivíduos. O que se percebe é que há um aumento substancial de 10% para 14% na redução da probabilidade de comprar CDs.

Tabela 6 - Estimativas de MQO e Variáveis Instrumentais - Probabilidade de comprar CDs e ir a Shows

Método	MQO	MQO	GMM	GMM	GMM	GMM	GMM	GMM	GMM
	Comprou CD					Foi a Show			Comprou CD Pirata
Variáveis	(1) ²	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Utiliza a Internet para Baixar Música	-0.0994*** (0.0168)	-0.141*** (0.0167)	-0.256** (0.116)	-0.370*** (0.119)	-0.338*** (0.120)	0.352*** (0.117)	0.293*** (0.108)	0.396*** (0.114)	-0.218** (0.0965)
Comprou CD Pirata	0.0963*** (0.0155)	0.107*** (0.0154)		-0.426*** (0.0783)	-0.374*** (0.0895)		0.0768 (0.0724)	0.242*** (0.0873)	
Constante	0.373*** (0.0483)	0.500*** (0.0485)	0.660*** (0.149)	0.866*** (0.149)	0.818*** (0.152)	0.587*** (0.150)	0.649*** (0.136)	0.494*** (0.145)	0.425*** (0.124)
Especificação de 1º estágio - (tabela 5)			(1)	(2), (4)	(3), (5)	(1)	(2), (4)	(3), (5)	(1)
Teste de sobreidentificação (p-valor) ¹	-	-	0.3339	0.4621	0.5003	0.3071	0.0044	0.4663	0.3613
Observações	7003	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.063	0.095	0.082	-0.099	-0.062	0.054	0.068	0.021	-0.009

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹H0: restrições de sobreidentificação são válidas

² A especificação da coluna (1) não controla o interesse declarado por música.

Essa mudança de magnitude nos coeficientes sugere que à medida que controlamos parte do gosto por música dos indivíduos as magnitudes dos coeficientes aumentam e mostram como as estimativas de *downloads* por MQO podem ser viesadas.

³⁴ Em relação aos métodos de estimação de variáveis dependentes limitadas, Probit e Logit, o uso do método de variáveis instrumentais é mais simples em termos computacionais e por ser uma abordagem menos estrutural requer menos hipóteses em termos de distribuições estatísticas. Ademais a única hipótese relevante para as estimativas por variáveis instrumentais, além da exogeneidade dos instrumentos, é que condicional as variáveis explicativas do modelo o valor esperado do erro seja igual a zero.

Já com relação ao consumo de CDs piratas, as estimativas apresentadas nas colunas (1) e (2) apresentam um acréscimo de 10% sobre a probabilidade de comprar CDs, um resultado que, em termos intuitivos, é inesperado. Não obstante este resultado se deve, em grande parte, a outros fatores não observáveis³⁵ além do gosto por música que podem influir nas decisões de consumo de CDs e CDs piratas.

Uma maneira que utilizamos para melhor controlar a influência de efeitos não observáveis é uso de variáveis instrumentais. As colunas (3) e (6) da tabela acima mostram estimativas dos efeitos de *download* sobre a probabilidade de comprar CDs e ir a shows quando não controlamos o consumo de CDs piratas. As estimativas foram geradas utilizando as estimativas de primeiro estágio referentes à coluna (1) da tabela 5. Os resultados gerados, embora viesados devido à omissão de pirataria, já mostram sinais compatíveis com os obtidos na seção de estatísticas descritivas. Observa-se que os indivíduos que declaram utilizar a Internet para baixar músicas possuem uma menor probabilidade de comprar CDs e uma maior probabilidade de ir a shows de música. Os coeficientes apresentam uma redução de cerca de 25% sobre a probabilidade de comprar CDs e um ganho de até 35% sobre a probabilidade de ir a shows de música.

Para cada especificação é fornecido o valor-p do teste de sobreidentificação na equação estrutural. Em todas as especificações o que se observa são valores de estatísticas que não permitem a rejeição da hipótese nula de sobreidentificação. Não obstante, é válido ressaltar que pelo modelo teórico se houver uma relação entre as escolhas de CDs piratas e *downloads* de música os resultados acima deixam de ser válidos.

Na coluna (9) da tabela 6 estimamos o efeito dos *downloads* sobre a probabilidade de ter comprado um CD pirata e verifica-se que existe uma relação de substituição entre *downloads* e consumo de CDs piratas. Vimos no modelo teórico que se houver uma relação na mesma direção entre CDs e CDs piratas, então, o viés gerado quando omitimos pirataria na equação estrutural atua no sentido de sobreestimar os coeficientes.

³⁵ A princípio poderia se pensar que a decisão sobre consumo de CDs de alguns indivíduos possa variar com o grau de fanatismo em relação a cada artista. Ou seja, para certos artistas que o indivíduo admira ele estaria disposto a pagar pelo CD original ao passo que para outros artistas, cujo consumo do álbum provê uma utilidade menor, ele compraria somente o CD pirata. Por outro lado, esse coeficiente positivo pode ser associado à minimização de risco por parte do consumidor. Pessoas que gostam de música podem, inicialmente, preferir comprar um CD pirata de um artista desconhecido e posteriormente comprar o CD original dado que as músicas desse artista se revelem boas o suficiente às exigências do consumidor.

Nas colunas (4), (5), (7) e (8) observamos justamente este tipo de comportamento. Quando controlamos o consumo de CDs piratas na equação estrutural observamos que há um aumento na magnitude dos coeficientes de *download* que explicam o consumo de CDs (colunas (4) e (5)) e uma redução nos coeficientes que explicam shows de música (coluna (7)). Outro ponto notável é a reversão do sinal do coeficiente de pirataria que em relação às estimativas de mínimos quadrados ordinários se torna negativo quando explica o consumo de CDs. Observa-se que há uma redução de até 42% sobre a probabilidade de consumo de CDs originais.

Importante ressaltar que as estimativas da coluna (8) e (5) excluem de suas estimativas de primeiro estágio a dummie “*é contra pirataria*”. Observa-se que quando excluímos essa variável do primeiro estágio há um aumento nos valores das probabilidades dos testes de sobreidentificação quando comparado às estimativas que incluem essa variável no primeiro estágio. Intuitivamente, essa variável, apesar de captar variações na tolerância a ilegalidade dos indivíduos, pode captar também algum fator que se associe ao respeito ou admiração por artistas que influi tanto no consumo de CDs, shows, *downloads* e CDs piratas. Não obstante, o que tabela 7 mostra é que esse fator é mais relevante nas escolhas de shows, cujas estimativas mostram uma maior sensibilidade a essa alteração na especificação de primeiro estágio. Nota-se que o coeficiente de pirataria torna-se significativo após a exclusão de aversão à pirataria do primeiro estágio o que implica em um ganho de até 24% sobre a probabilidade de ir a shows de música.

Os resultados acima corroboram com as suspeitas levantadas na seção de estatísticas descritivas. Embora, haja uma redução substancial na probabilidade de comprar CDs devido ao ato de baixar músicas pela Internet, esse efeito deletério é, em parte, compensado por um efeito positivo na probabilidade de ir a shows de música. Isto é uma evidência de que não necessariamente a livre circulação de conteúdo protegido na Internet inibe os incentivos de criação como aponta a literatura teórica de direitos de propriedade.

Outro resultado que reforça esta hipótese surge quando estimamos o efeito de fazer *download* de músicas sobre os gastos com shows de música.

Tabela 7 - Efeito dos Downloads e CDs Piratas Sobre o Gasto com Shows - GMM

Variáveis Explicativas	Gastos em Reais com Shows de Música							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Utiliza a Internet para baixar músicas	167.2*** (42.85)				177.8*** (45.89)			
Baixou CD		114.5*** (28.35)				122.9*** (30.55)		
Baixou Música			168.1*** (43.29)				192.9*** (49.18)	
Faz download frequentemente ¹				141.8*** (37.98)				159.8*** (42.60)
Comprou CD pirata	-23.75 (36.67)	-22.16 (36.00)	-56.76 (41.60)	-38.76 (38.41)				
Teste de sobreidentificação (p-valor) ²	0.3132	0.2875	0.4775	0.2184	0.3605	0.3444	0.6650	0.3007
Observações	6940	6940	6940	6940	6940	6940	6940	6940
R ²	-0.001	0.034	-0.084	-0.002	-0.007	0.030	-0.120	-0.018

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Faz download de músicas pelo menos 1 vez na semana²H0: restrições de sobreidentificação são válidas

Nota: As estimativas acima são computadas utilizando as especificações de primeiro estágio das colunas (3) e (5) da tabela 5.

Utilizando a especificação de primeiro estágio das colunas (3) e (5) da tabela 5, os resultados da tabela acima mostram que as pessoas que fazem *downloads* de música têm um gasto com shows superior em até cerca de 190,00. Na amostra utilizada não possuímos informações acerca do gasto com CDs de música. Entretanto, podemos estimar o efeito de baixar músicas sobre a quantidade de CDs adquiridos nos últimos 6 meses. Devido uma censura observada na quantidade de CDs adquiridos utilizamos o método de estimação Tobit com variáveis instrumentais para levar em conta a censura presente nos dados referentes ao consumo de CDs.

Tabela 8 - Efeitos de Downloads sobre Quantidade de CDs - TOBIT

Variáveis	CDs - Quantidade Consumida	
	(1)	(2)
Utiliza a Internet para Baixar Músicas	-9.94* (5.321)	-14.71*** (5.468)
Efeito Marginal de Download (dy/dx)	-2.75* (1.670)	-4.07*** (1.72)
Comprou CDs piratas		-19.34*** (4.428)
Efeito Marginal de Pirataria(dy/dx)		-5.36*** (1.390)
Constante	1.17 (6.798)	9.802 (6.940)
Observações	7003	7003

Desvios Padrões entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal,

Anos de Educação na USP e interesse declarado por música

Instrumentos : Desfragmentou Disco, Formatou Computador

Utiliza Antivirus, Conhece Memória Ram, Conhece Memória HD,

Conhece Processador e N° de Produtos de Feira ou Camelô

A tabela acima mostra que as pessoas que utilizam a Internet para baixar músicas deixam de comprar cerca de 4 CDs em relação às pessoas que não baixam. Essas estimativas permitem a geração de um comparativo em relação às estimativas de gastos com shows. Segundo o relatório³⁶ da ABPD no ano de 2009 a indústria fonográfica brasileira vendeu cerca 20,3 milhões de CDs e obteve uma receita de aproximadamente R\$ 215 milhões de reais. Utilizando esses números vemos que o preço médio do CD no mercado nacional é de cerca de R\$ 10,50. Por meio dos coeficientes estimados acima vemos que uso da Internet para a finalidade de *download* de músicas impõe um ganho médio de 180 reais por indivíduo com shows e uma perda de até R\$ 42,70 devido à substituição de CDs por *downloads*. Ou seja, quando levamos em conta apenas os impactos dos *downloads* de música a Internet gera um ganho líquido ao mercado de aproximadamente R\$ 137,26 por indivíduo que faz *download* de músicas. Supondo que a demanda por CDs seja preço inelástica, o preço médio do CD que anularia esse excedente seria de aproximadamente R\$ 48,00.

Este efeito positivo dos *downloads* sobre shows é uma evidência que confirma os resultados de Mortimer e Sorensen (2005) que atestam que os *downloads* de música têm um efeito positivo sobre a criação de música via efeito preço. Como fazer *download* de música pode implicar num aumento da demanda por shows, os artistas têm maiores incentivos a fazerem mais shows dado que esse aumento na demanda impulsiona os preços dos ingressos.

No apêndice C, as tabelas C.1 e C.2 provem estimativas adicionais dos efeitos de *download* e consumo de CDs piratas sobre a probabilidade de ir a Shows. Nestas tabelas são utilizadas medidas alternativas de atividade de *download* que indicam se o indivíduo baixou pelo menos um CD na Internet, uma música avulsa e se ele utiliza a Internet pelo menos uma vez na semana para baixar músicas. O resultado que se observa quando utilizamos essas demais medidas é consistente com os resultados observados na tabela 7 desta seção.

O resultado negativo sobre o consumo de CDs também é consistente com o resultado gerado por uma pesquisa prévia cujos resultados são mostrados no apêndice B. Nesta pesquisa encontramos efeitos negativos significantes dos *downloads* sobre o consumo de CDs e efeitos positivos sobre o consumo de bens subordinados. Não obstante, estes resultados foram

³⁶ Disponível em: http://www.abpd.org.br/downloads/Final_Publicacao_09_2010_CB.pdf

gerados com o uso de instrumentos que possivelmente se correlacionam com o gosto por música. Uma evidência são os coeficientes estimados na tabela C.4 (ver apêndice C) que utilizam a base de dados construída por Internet e mostram efeitos insignificantes dos *downloads* sobre o consumo de bens subordinados³⁷. Isso sugere que o resultado positivo sobre bens subordinados encontrado quando é utilizada a amostra coletada em salas de aula possivelmente é devido a um viés gerado pela correlação dos instrumentos com o gosto por música.

Por fim, é importante ressaltar que quando utilizamos as contagens de consumo de CDs, *Downloads*, CDs piratas e Shows para estimar a equação (7) obtemos estimativas que são estatisticamente insignificantes para explicar o consumo de CDs. Este resultado é reportado na tabela C.5 do apêndice C. Se ignorarmos a possibilidade desse resultado ser devido ao erro de medida presente nas variáveis, o que os resultados acima mostram é que baixar uma música ou um CD a mais tem um efeito pouco expressivo sobre a quantidade de CDs consumida. Quando se utiliza variáveis dicotômicas para explicar a probabilidade de baixar músicas o que se percebe é que o efeito dos *downloads* sobre CDs é mais uma questão de um efeito fixo do que marginal.

³⁷ A tabela também mostra que existem efeitos positivos significantes a um nível de significância de 10% do consumo de CDs piratas sobre o consumo de produtos subordinados. Este efeito positivo pode ser devido ao fato de que camisetas de bandas e ou outros produtos com imagens de artistas serem amplamente ofertados em mercados informais onde os CDs piratas também são ofertados.

7. ANÁLISE DE ROBUSTEZ

De forma a certificar os resultados apresentados na seção anterior será realizado nesta seção alguns exercícios que buscam investigar a sensibilidade dos resultados gerados a mudanças nas especificações das equações de primeiro estágio e também a restrições na amostra com base em algumas características individuais.

7.1 Mudando Especificações

Vimos na discussão dos instrumentos que um grupo de variáveis instrumentais é influenciado pelas *dummies* de interesse declarado por música ao passo que outro grupo é influenciado pela *dumme* “*toca instrumento musical*”. Por sua vez, essas evidências mostram que cada conjunto de instrumentos pode ser correlacionado com fatores específicos do gosto por música não observado. Utilizados conjuntamente, os instrumentos podem levar em conta a influência de cada fator sobre as estimativas. O exercício empírico a seguir consiste em utilizar diferentes combinações de instrumentos no primeiro estágio e verificar como os coeficientes da equação estrutural respondem a essas mudanças. Na tabela abaixo fornecemos as estimativas de primeiro estágio das especificações consideradas.

Tabela 9 - Estimativas de Primeiro Estágio – Análise de Robustez - MQO

Variáveis Instrumentais	Utiliza a Internet para Baixar Músicas							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Desfragmentou Computador	0.0490*** (0.0104)	0.0511*** (0.0104)			0.0436*** (0.0103)			
Formatou Computador	0.0373*** (0.00915)	0.0348*** (0.00916)			0.0268*** (0.00923)			
Utiliza Antivirus	0.0542*** (0.0174)	0.0521*** (0.0173)						
É contra Pirataria	-0.0513*** (0.00818)		-0.0536*** (0.00818)			-0.0464*** (0.00841)	-0.0532*** (0.00824)	
Nº de Produtos de Feira ou Camelô		0.0108*** (0.00227)		0.0116*** (0.00228)	0.0110*** (0.00227)	0.00838*** (0.00233)		0.0118*** (0.00229)
Conhece a Memória Ram			0.0130 (0.0174)	0.0143 (0.0174)				
Conhece a Memória de Disco Rígido			0.0857*** (0.0191)	0.0844*** (0.0191)	0.0784*** (0.0118)			
Conhece o Processador			0.00805 (0.0150)	0.00647 (0.0151)				
Observações	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.154	0.151	0.154	0.151	0.156	0.143	0.141	0.139

Desvios Padrões robustos entre Parênteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP

e Interesse declarado por música

As especificações da coluna (1) e (2) consideram somente as *dummies* de atividades de informática ao passo que as especificações da coluna (3) e (4) apenas as *dummies* de itens da configuração do computador. A coluna (5) exclui da especificação as variáveis que não se mostram significantes na especificação da coluna (3) da tabela 7 e por fim, as colunas (6), (7) e (8) utilizam instrumentos ligados apenas ao consumo de CDs piratas.

Os resultados obtidos quando utilizamos estas especificações são mostrados na tabela abaixo. Em relação às estimativas que levam em consideração itens da configuração no primeiro estágio (colunas (3), (4), (8) e (9)), as estimativas que utilizam as *dummies* de atividades de informática (colunas (1), (2), (6) e (7)) mostram coeficientes que são menores quando explicam a probabilidade de comprar CDs e maiores quando explicam a probabilidade de ir a shows. Não obstante, os sinais obtidos com todas as especificações são preservados. As reduções na probabilidade de compra CDs devido aos *downloads* de música variam de um mínimo de 29% a um máximo de 45% ao passo que os ganhos sobre a probabilidade de ir a shows oscilam no intervalo de 31% a 42%.

Tabela 10 - Análise de Robustês - Equação Estrutural - GMM

Variáveis	Comprou CD					Foi a Show			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Utiliza a Internet para Baixar Músicas	-0.300** (0.145)	-0.292** (0.143)	-0.455*** (0.144)	-0.432*** (0.143)	-0.390*** (0.122)	0.341*** (0.132)	0.419*** (0.138)	0.224* (0.128)	0.312** (0.135)
Comprou CD pirata	-0.530*** (0.104)	-0.401*** (0.0931)	-0.446*** (0.105)	-0.347*** (0.0929)	-0.370*** (0.0912)	-0.148 (0.0960)	0.225** (0.0902)	-0.0862 (0.0951)	0.266*** (0.0899)
Especificação de 1º estágio - tabela 9	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)
Teste de sobreidentificação (p-valor) ²	0.4141	0.3392	0.5730	0.6890	0.6820	0.2111	0.2822	0.1398	0.3210
Observações	7003	7003	7003	7003	7003	7003	7003	7003	7003
R ²	-0.156	-0.068	-0.132	-0.067	-0.069	0.036	0.017	0.070	0.036

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Faz download de músicas pelo menos 1 vez na semana

²H0: restrições de sobreidentificação são válidas

Tabela 10 - Continuação

Variáveis	Foi a Show				
	(10)	(11)	(12)	(13)	(14)
Utiliza a Internet para Baixar Músicas	0.355*** (0.116)	0.344* (0.192)	-0.0450 (0.215)	1.154*** (0.348)	-3.264 (2.636)
Comprou CD pirata	0.247*** (0.0883)				1.319 (0.951)
Especificação de 1º estágio - tabela 9	(5)	(6)	(7)	(8)	(6)
Teste de sobreidentificação (p-valor) ²	0.3812	0.001	0.0000	0.000	0.000
Observações	7003	7003	7003	7003	7003
R ²	0.031	0.056	0.081	-0.454	-6.059

Comparando essas estimativas com as obtidas na tabela 6 vemos que as magnitudes dos coeficientes das estimativas da tabela 6 se encontram entre os obtidos na tabela acima. Mesmo havendo uma mudança na magnitude dos coeficientes os resultados são qualitativamente consistentes. Já em relação aos coeficientes relacionados à pirataria de CDs nota-se uma mudança sutil em relação aos coeficientes da tabela 7. Isso se deve, em grande parte, a pouca sensibilidade do consumo de pirataria aos instrumentos utilizados para *downloads* de música que se ligam a habilidade de informática.

As colunas (11), (12), (13) e (14) estimam os efeitos dos *downloads* e CDs piratas sobre shows utilizando no primeiro estágio apenas variáveis ligadas ao consumo de CDs piratas. Nas colunas (11) a (13) devido à omissão de pirataria na equação estrutural gera-se um problema de identificação, pois os instrumentos utilizados são altamente correlacionados com o consumo de pirataria. Não obstante, mesmo com esse problema de identificação, nas colunas (11) e (13) consegue-se obter coeficientes que apontam na mesma direção das análises anteriores. Já na coluna (14), que leva em conta o consumo de pirataria, o que se observa são estimativas estatisticamente não significantes e uma reversão do sinal do coeficiente de *downloads* sobre shows. Essa reversão de sinal, em grande parte, é atribuída a uma multicolinearidade devida à relação de substituição entre *download* e pirataria física.

Como os instrumentos utilizados possuem um efeito marginal mais intenso sobre a probabilidade de comprar CDs piratas do que de fazer *downloads*, o efeito indireto de pirataria sobre *downloads*, que é negativo, amplifica o viés devido à multicolinearidade entre as duas variáveis. Por sua vez, isto pode alterar o sinal do coeficiente que é positivo em negativo.

Por fim, uma última alteração na estrutura do modelo estimado é checar como as estimativas na equação estrutural se alteram quando não levamos em conta alguns controles como os institutos em que os alunos estão matriculados, a quantidade de anos que estudam na universidade e a cor ou raça dos respondentes. A tabela abaixo estima o efeito dos *downloads* sobre a probabilidade de ir a shows e comprar CDs utilizando diferentes combinações desses controles mencionados anteriormente. Comparando os resultados da tabela abaixo com os da tabela 6 observamos que os coeficientes estimados são pouco influenciados pelo controle ou não dessas variáveis.

Tabela 11 - Alterando a Equação Estrutural - GMM

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variáveis	Foi a Show	Foi a Show	Foi a Show	Foi a Show	Comprou CD	Comprou CD	Comprou CD	Comprou CD
Utiliza a Internet para Fazer Download	0.339*** (0.115)	0.369*** (0.114)	0.389*** (0.113)	0.332*** (0.116)	-0.403*** (0.123)	-0.366*** (0.119)	-0.360*** (0.119)	-0.426*** (0.122)
Comprou CD pirata	0.220*** (0.0844)	0.272*** (0.0880)	0.233*** (0.0866)	0.253*** (0.0832)	-0.398*** (0.0879)	-0.385*** (0.0896)	-0.375*** (0.0898)	-0.381*** (0.0852)
Institutos	não	sim	sim	não	não	sim	sim	não
Anos de Estudo na USP	sim	sim	não	não	sim	não	sim	não
Cor	sim	não	sim	não	sim	não	não	não
Observações	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.028	0.013	0.024	0.012	-0.100	-0.073	-0.067	-0.100

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, renda familiar mensal e interesse declarado por música

Nota: As estimativas acima são computadas utilizando-se as especificações de primeiro estágio das colunas (3) e (5) da tabela 5

O instituto o qual indivíduo está matriculado poderia ter alguma influência sobre o perfil do aluno no sentido de que os cursos envolvem atividades de naturezas distintas. Um exemplo claro é o fato de um aluno ser matriculado na Escola de Comunicação e Artes - ECA e outro matriculado no Instituto de Matemática e Estatística – IME. É razoável pensar que o aluno matriculado na ECA tenha um maior interesse por entretenimento, artes e cultura do que o aluno matriculado no IME.

Já o tempo que o indivíduo estuda na universidade também poderia influenciar esse consumo devido a vários aspectos além dos anos de educação na universidade. Um desses aspectos é que quanto mais tempo um indivíduo convive na universidade mais pessoas ele conhece e essas pessoas podem de alguma forma afetar o consumo de música.

Por fim os exercícios acima mostram que as estimativas também são pouco afetadas pelo controle ou não da cor ou raça dos indivíduos. Isso, em parte, implica que o consumo de música dos indivíduos é pouco relacionado com fatores raciais ou étnicos.

7.2 Análise de Subamostras

Outro aspecto importante de nossa análise é saber a sensibilidade dos resultados quando restringimos a análise para um grupo de indivíduos com características específicas. No apêndice C provemos uma tabela³⁸ que mostra uma maior propensão para o consumo de CDs entre os indivíduos que escutam jazz, blues, soul e música clássica. Isso reflete uma maior intensidade pela preferência por CDs e um maior gosto por música. A princípio, por ser um grupo com uma maior propensão do consumo de CDs existe uma maior chance de se observar também um efeito dos *downloads* sobre o consumo de CDs que seja menos deletério ou mesmo positivo. Portanto, se mesmo para esse grupo específico encontrarmos um efeito substitutivo dos *downloads*, teremos um resultado que fortalece os argumentos providos anteriormente.

Outro ponto relevante é a influência do grau de informação dos indivíduos sobre as escolhas dos instrumentos ligados a habilidade em informática e shows de música. A princípio pode-se pensar que indivíduos com um maior grau de habilidade em informática sejam indivíduos mais informados e conseqüentemente vão mais a shows de música. No questionário algumas perguntas são relacionadas ao posicionamento dos indivíduos acerca de temas como legalização da maconha, direção alcoolizada e restrições ao porte de armas. Basicamente são questões que solicitam aos respondentes que se posicionem a respeito desses temas indicando se são a favor, contra, se não sabem ou não tem posição definida. Supondo que os indivíduos formam suas opiniões a partir de informações recebidas podemos utilizar o grupo de indivíduos que não tem posições definidas acerca dos temas acima para testar se há algum efeito informacional sobre as estimativas dos *downloads* sobre shows. Parte-se do pressuposto de que indivíduos que não tem posições definidas acerca desses temas são indivíduos que não se informaram ou não procuram se informar o suficiente para formar uma opinião definida.

³⁸ Ver a tabela C.3 no apêndice C.

Também utilizamos uma amostra restrita aos indivíduos que não possuem um partido político. Pelos mesmos motivos explicitados anteriormente indivíduos não-politizados podem de certa forma se revelar indivíduos menos informados. Na outra ponta, utiliza-se também uma amostra restrita a indivíduos que não possuem laptop. Indivíduos que possuem laptop têm uma maior possibilidade de se conectar a Internet no campus universitário e com isso se informarem melhor em relação a indivíduos que só podem acessar a Internet por meio dos laboratórios. A tabela abaixo apresenta os resultados de tal exercício empírico.

Tabela 12 - Efeito dos *Downloads* sobre Shows e CDs – Análise de Subamostras - GMM

	(1)	(2)	(3)	(4)	(5)	(7)	(6)	(8)	(9)
Variáveis	Comprou CD	Foi a Show	Gastos com Shows	Foi a Show	Gastos com Shows	Foi a Show	Gastos com Shows	Foi a Show	Gastos com Shows
Utiliza a Internet para Fazer Download	-0.544*** (0.180)	0.0972 (0.150)	211.1*** (63.59)	0.547*** (0.185)	177.2** (69.84)	0.441** (0.223)	168.8** (73.65)	0.440** (0.202)	191.1*** (57.26)
Comprou CD pirata	-0.331** (0.161)	0.235* (0.140)	-71.85 (58.61)	0.321* (0.180)	-33.02 (78.75)	0.354** (0.155)	29.09 (70.93)	0.552*** (0.180)	56.59 (68.42)
Observações	Grupo1 3297	Grupo1 3297	Grupo1 3246	Grupo2 3047	Grupo2 3004	Grupo3 1747	Grupo3 1721	Grupo4 2062	Grupo4 2032
R ²	-0.063	0.076	-0.030	0.005	0.004	0.072	-0.009	-0.024	-0.003

Desvios Padrões robustos entre Parênteses

*** p<0.01, ** p<0.05, * p<0

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

Grupo 1 : Pessoas que escutam Jazz, Soul, Blues e Música Clássica

Grupo 2: Pessoas que não possuem um partido político

Grupo 3: Pessoas que não possuem um Laptop

Grupo 4: Pessoas que não possuem posições definidas acerca de temas como: Legalização da Maconha

Direção alcoolizada, Posição Ideológica e Restrições ao Porte de Armas.

Nota: As estimativas acima são computadas utilizando-se as especificações de primeiro estágio da coluna (3) e (5) da tabela 5

Utilizando as especificações de primeiro estágio das colunas (3) e (5) da tabela 5 é estimado o efeito dos *downloads* sobre a probabilidade de comprar CDs, ir a shows e gastos com Shows de música. Quando restringimos a amostra ao primeiro grupo observamos que não há uma inversão do sinal do coeficiente de *downloads* sobre CDs. Observa-se que o efeito negativo sobre a probabilidade de comprar CDs persiste. Para esse grupo não se observa o efeito positivo dos *downloads* sobre a probabilidade de ir shows. Não obstante, os gastos com shows são significantes, esses coeficientes mostram que, para este grupo, as pessoas que baixam música têm a mesma probabilidade de ir a shows em relação aos que não baixam, mas os gastos são mais acentuados.

As demais estimativas apontam que mesmo restringindo a amostra a grupos de indivíduos que possivelmente são pouco informados os resultados positivos de *downloads* sobre shows são mantidos. Por sua vez, isso fornece uma evidência contra a hipótese de que poderia haver

algum efeito informacional sobre as estimativas fornecidas anteriormente. Se houvesse tal efeito deveríamos observar um decréscimo tanto na magnitude dos coeficientes quanto em sua significância, algo que não é observado na tabela acima.

8.DISSCUSSÃO E CONCLUSÃO

A indústria fonográfica Brasileira observou uma queda de cerca de 70% sobre as receitas com produtos de áudio entre os anos de 1999 e 2007. Comparado com o declínio nas receitas observado na indústria fonográfica mundial, no Brasil essa queda teve uma maior intensidade. Este fenômeno pode ser devido a características institucionais presentes no Brasil que podem ter influência sobre o comportamento de *download* dos brasileiros.

Nesse trabalho nós utilizamos uma base de dados com informações de 7.147 estudantes da Universidade de São Paulo para medir o impacto dos *downloads* de música sobre o consumo de CDs e Shows de música. Os resultados obtidos mostram evidências de que os *downloads* reduzem a probabilidade de comprar CDs. Este resultado também é confirmado por meio do uso de uma amostra secundária que foi obtida por meio da aplicação de questionários em salas de aula da USP. Não obstante, esse efeito negativo é ao menos parcialmente compensado por um efeito positivo dos *downloads* sobre a demanda por Shows de música. Utilizando medidas de habilidade em informática e tolerância a ilegalidade nós estimamos efeitos em que os *downloads* reduzem em até 45% a probabilidade de comprar CDs ao passo que produzem um aumento de até 40% sobre a probabilidade de ir a shows. Essa última estimativa é confirmada com outro resultado que mostra que os indivíduos da amostra que fazem *download* têm um gasto médio com shows de música superior em até 200 reais.

Testes adicionais de robustez que levam em conta mudanças nas especificações de primeiro estágio e na composição da amostra utilizada não mostram alterações bruscas em relação aos resultados obtidos na seção de abordagem empírica e reforçam os resultados obtidos na análise principal deste trabalho.

Quando comparamos nossos resultados com os obtidos por Zentner (2006) observamos que os coeficientes estimados são compatíveis com a queda mais acentuada observada nas receitas da indústria fonográfica Brasileira. Em parte, a queda nas receitas de produtos de áudio pode ser devido ao maior efeito redutivo que os *downloads* têm sobre a probabilidade de comprar CDs. Não obstante, os resultados acima, especialmente os referentes ao impacto dos *downloads* sobre Shows de música, mostram que os *downloads* de música, podem ter um efeito positivo

sobre os incentivos a criação de música. Ademais, nós também mostramos que existe uma relação de substituição entre *downloads* e consumo de CDs piratas.

Por fim, nossa última contribuição é mostrar que ao se tentar estimar os efeitos dos *downloads* sobre o consumo de CDs é necessário levar em conta o consumo de CDs piratas. As previsões feitas pelo modelo teórico de que, quando há uma relação entre *downloads* e consumo de CDs piratas, a omissão de pirataria pode subestimar os coeficientes são confirmadas pelas estimativas apresentadas. Em geral, quando controlamos o consumo de CDs piratas na equação estrutural o que observamos é um aumento da magnitude dos efeitos dos *downloads* sobre o consumo de CDs e uma redução dos efeitos dos *downloads* sobre shows de música.

REFERÊNCIAS

ABPD. *O Mercado Brasileiro de Música 2008*. Rio de Janeiro. 2008

ASVANUND, A, K et al. *An Empirical Analysis of Network Externalities in Peer to Peer Music-Sharing Networks*. **Information Systems Research**, Vol.15, p. 155-174, 2004

BAKOS, Yannis, Erik BRYNJOLFSSON, Douglas LICHTMAN. *Shared Information Goods*. **Journal of Law and Economics**. Vol.42, p. 117–56.

BHATTACHARJEE, S et al. *Impact of Legal Threats on Online Music Sharing Activity: An Analysis of Music Industry Legal Actions*. **Journal of Political Economy**, vol. 59 , 2006

BHATTACHARJEE, S et al. *Whatever happened to payola? An empirical analysis of online music sharing*. **Decision Support Systems**, Vol. 42, p. 104-120, 2006

BLACKBURN, D. *On-line Piracy and Recorded Music Sales*. **Working Paper**, 2004

BOUNIE, D, M, BUREAU, P, WAELBROECK. *Pirates or Explores? Analysis of Music Consumption in French Graduate Schools*. **Working Paper**

CAMERON, A. Colin & Pravin K. TRIVEDI. **Microeconometrics Using Stata**. Stata Press. 2009

CORTEZ, Igor Siqueira. *Pirataria de Música na Internet: Uma Análise de Bem-Estar Social. Trabalho de Conclusão de Curso*, Universidade de Brasília, 2007

CURIEN, N et al. *Towards a New Business Model for the Music Industry: Accomodating Piracy through Ancillary Products*. **Laboratoire d'économetrie**, Conservatoire national des Arts et Métiers, 2004

GAYER, A & O. SHY. *Internet and Peer-to-Peer Distributions in Markets for Digital Products*. **Economics Letters**, Vol. 81,p. 51-57, 2003

GAYER, A & O. SHY. *Copyright Protection and Hardware Taxation*. **Information Economics & Policy**, Vol.15, p.467-483, 2003

GENTZKOW, Matthew. *Valuing New Goods in model with complementarity: Online Newspaper*. **The American Economic Review**, Vol.97, No.3, 2007

HAYASHI, Fumio. **Econometrics**. Princeton University Press, 2000

JOHNSON,W. *The Economics of Copying*. **Journal of Political Economy**. Vol.93, p.158-174, 1985

MORTIMER, J. H & A. SORENSEN. *Supply Responses to Digital Distribution: Recorded Music and Live Performances*. **Working Paper**. Harvard University, 2005

NOVOS, I.E & M. WALDMAN. *The Effects of Increased Copyright Protection: An Analytic Approach*, ***Journal of Political Economy***, Vol.92, p.236-246, 1984

OBERHOLZER F. & K. STRUMPF. *The Effect of File Sharing on Record Sales: An Empirical Analysis*. ***Journal of Political Economy***, vol.115, no. 1, 2007

OBERHOLZER F. & K. STRUMPF. *File Sharing and Copyright* . ***Working Paper***, 2009

ROB, R. & J WALDFOGEL. *Piracy on the High C'S: Music Downloading, Sales Displacement, and Social Welfare in Sample of College Students*, ***Journal of Law and Economics***, Vol.49, 2006

WALDFOGEL, J. *Music File Sharing and Sales Displacement in The iTunes Era*. ***Working Paper***. Universty of Pennsylvania, 2009

WOOLDRIDGE, J. ***Econometric Analysis of Cross Section and Panel Data***. MIT Press, 2002

ZENTNER, A. *Measuring The Effect of File Sharing on Music Purchases*. ***Journal of Law and Economics***, Vol 49 , 2006

APÊNDICES

Apêndice A – Estatísticas descritivas complementares

Apêndice B – Resultados empíricos obtidos com a amostra coletada em salas de aula

Apêndice C – Tabelas complementares ao capítulo 6 e 7

Apêndice D – Questionário

A - ESTATÍSTICAS DESCRITIVAS COMPLEMENTARES

A.1 - Amostra coletada em Salas de Aula – Dados e Estatísticas Descritivas

A coleta dos dados desta amostragem foi realizada entre os meses de março e maio de 2009. Foi selecionada uma lista de professores que ministram aulas de matérias obrigatórias do primeiro e segundo ano de graduação da USP e, em seguida, foi enviado um e-mail solicitando a aplicação do questionário nos alunos das respectivas turmas. Sendo assim, o questionário foi aplicado nas turmas em que houve autorização dos respectivos professores³⁹. Neste processo foi coletada uma amostra de 716 alunos da USP dos quais a maioria são alunos matriculados nos anos de 2008 e 2009 de 9 institutos da cidade universitária. No apêndice A.2 uma tabela fornece a distribuição dos institutos entre as amostras analisadas. As principais estatísticas descritivas são mostradas na tabela abaixo.

É interessante notar que em relação à amostra feita pela Internet, a idade média observada é menor devido aos alunos, em maioria, serem de primeiro e segundo ano da graduação. Dois outros pontos importantes são a presença de uma menor parcela de pessoas que possuem laptop e uma maior de pessoas que consomem CDs piratas. Esta última diferença se deve a uma alteração feita no questionário utilizado na pesquisa *on-line*. Basicamente, ao invés de se perguntar a coleção de CDs piratas perguntou-se o consumo para o período de 6 meses.

Outro ponto relevante é que assim como na amostra *on-line* existe uma maior proporção de pessoas que comprou CDs no extrato de pessoas que não fazem *download* e uma maior proporção de pessoas que foi a shows entre as pessoas que fazem *download*.

³⁹ Um padrão observado foi que os professores de institutos de áreas exatas como matemática, física, e estatística se mostraram relutantes em ceder um horário para aplicar os questionários ao passo que os professores de áreas humanas como psicologia, economia, história e geografia mostraram uma maior aceitação à aplicação do questionário

Tabela A.1 - Estatísticas Sumárias da Amostra Coletada em Salas de Aula

<i>Estatísticas</i>	
<i>Amostra Total</i>	N=716
<i>idade média</i>	20.8
	%
Pessoas do Sexo Masculino	61.0
Pessoas que baixam música pela internet	87.8
Pessoas com conexão à internet em casa	93.4
Possuem Laptop com conexão wireless	39.2
Pessoas que compraram CD nos últimos 6 meses	36.5
Pessoas que foram a Shows nos últimos 6 meses	56.3
Pessoas que compraram Bens subordinados	33.1
Pessoas com conexão à internet (banda-larga)	88.1
Pessoas que compram CDs piratas	82.8
<i>Pessoas que fazem download de músicas</i>	N=628
	%
Pessoas que compraram CDs nos últimos 6 meses	34.2
Pessoas que não compraram CDs nos últimos 6 meses	65.8
Pessoas com conexão à internet (banda-larga)	91.7
Pessoas que consomem CD pirata	83.9
Pessoas que foram a Shows nos últimos 6 meses	57.8
<i>Pessoas que não fazem download de músicas</i>	N=87
	%
Pessoas que compraram CDs nos últimos 6 meses	52.8
Pessoas que não compraram CDs nos últimos 6 meses	47.2
Pessoas com conexão à internet (banda-larga)	65.5
Pessoas que consomem CD pirata	74.7
Pessoas que foram a Shows nos últimos 6 meses	46.0

A.2 - Institutos observados por tipo de amostragem.

Os institutos que compõem as duas amostras analisadas são descritos na tabela, abaixo. Dentre elas, a amostragem realizada por Internet é a mais abrangente e a que melhor representa as características da população de estudantes da Universidade de São Paulo. Um padrão que se observa nas duas amostragens é que os alunos das unidades FEA, FFLCH e POLI são maioria.

Para cada instituto e em cada amostra foi calculado a porcentagem das pessoas que declarou ter comprado um CD original e a porcentagem de pessoas que baixou música nos últimos 6 meses. Quando calculamos a média das porcentagens nas duas amostras, observamos que a proporção média de pessoas que afirmam ter comprado CDs entre os institutos é de cerca de um terço e a proporção média das pessoas que afirmam ter baixado música aproxima-se de 73% entre os institutos.

Tabela A.2 - Distribuição dos Institutos entre as Amostras

Instituto	Amostra em Sala de Aula				Amostra On-line			
	Frequência	%	Comprou CDS(%)	Baixou Músicas (%)	Frequência	%	Comprou CDS(%)	Baixou Músicas (%)
EACH					386	5,4	36,0	73,6
ECA					212	3,0	42,9	75,9
EE					160	2,2	30,0	62,5
EEFE					55	0,8	21,8	78,2
ESALQ					349	4,9	25,8	72,5
FAU					147	2,1	36,1	76,2
FCF					149	2,1	32,9	65,8
FD					156	2,2	49,4	75,0
FE					142	2,0	57,0	64,8
FEA	262,0	36,6	36,6	75,6	637	8,9	31,2	73,9
FFLCH	219,0	30,6	40,6	75,3	986	13,8	42,3	75,0
FM					287	4,0	41,5	64,1
FMVZ					77	1,1	48,1	67,5
FO					115	1,6	32,2	66,1
FSP					155	2,2	36,8	63,9
FZEA					108	1,5	25,9	81,5
IAG					56	0,8	37,5	78,6
IB	26,0	3,6	38,5	65,4	136	1,9	30,9	72,8
ICB					73	1,0	31,5	72,6
ICMC					192	2,7	21,4	79,7
IF	1,0	0,1	0	100	217	3,0	31,8	71,0
IGc					57	0,8	33,3	73,7
IME	63,0	8,8	39,7	85,7	181	2,5	35,9	79,0
IO	1,0	0,1	0,0		43	0,6	25,6	72,1
IP	55,0	7,7	38,2	63,6	100	1,4	37,0	69,0
IQ	1,0	0,1	0	100	118	1,7	35,6	72,0
IRI					29	0,4	37,9	75,9
POLI	88,0	12,3	23,9	73,9	551	7,7	31,6	77,9
Outros					1273	17,8	28,6	75,1
Média			31,1	73,3			34,8	72,6
Total	716,0	100			7147	100		

B - Resultados da Análise Empírica feita com Amostra coletada em Salas de Aula

A tabela abaixo provê os resultados obtidos quando utilizamos os dados referentes à amostra coletada em salas de aula. A coluna (1) mostra o coeficiente obtido quando utilizamos uma regressão simples para estimarmos o efeito dos *downloads* sobre o consumo de CDs. Devido à endogeneidade da variável de *download* obtemos um coeficiente positivo e não significativo. Entretanto, quando utilizamos o método de variáveis instrumentais observamos uma reversão do sinal. As colunas (2) e (3) apresentam um resultado negativo e significativo ao um nível de significância de 5%. Os coeficientes indicam que as pessoas que baixam CDs pela Internet compraram 3,5 CDs a menos do que as pessoas que não baixam.

Tabela B.1 Resultados - Amostra Coletada Em Salas de Aula

MÉTODO	(1) OLS	(2) GMM	(3) GMM	(4) GMM	(5) GMM	(6) GMM	(7) GMM
VARIÁVEIS	CDs	CDs	CDs	Comprou CD	Comprou Bem Subordinado	Comprou Bem Subordinado	Foi a Show
Baixou CD	0.299 (0.347)	-3.349** (1.470)	-3.598** (1.494)			0.343*** (0.129)	-0.047 (0.136)
Faz download pelo menos 1x/semana				-0.370* (0.211)	0.493** (0.193)		
Teste de sobreidentificação (p-valor) ¹		0.2366	0.8745	0.4078	0.1380	0.0729	0.0740
Observações	502	502	502	502	502	502	502
R ²	0.047	0.019	0.011	0.041	0.040	0.089	0.0226
Especificação de primeiro estágio	-	(1)	(2)	(4)	(4)	(2)	(2)
Desvios Padrões robustos entre Parenteses							
*** p<0.01, ** p<0.05, * p<0.1							
Controles: idade, sexo, renda familiar mensal, Anos de Educação na USP e interesse declarado por música							
¹ H0: restrições de sobreidentificação são válidas							

Na coluna (4) utilizamos uma medida sobre a frequência com que a pessoa faz *download*. A idéia é saber se as pessoas com um comportamento mais freqüente de *download* também tem uma menor probabilidade de baixar músicas. O coeficiente estimado mostra que as pessoas que baixam músicas pelo menos uma vez na semana têm uma probabilidade de ter comprado CDs 37% menor do que as pessoas que não baixam músicas freqüentemente.

Na outra ponta, quando olhamos para o consumo de bens subordinados observamos um efeito positivo dos *downloads*. As pessoas que baixam CDs pela Internet têm uma probabilidade 34% maior de ter comprado produtos associados ao artista ou banda do que as pessoas que não baixam músicas. Já com relação aos shows, utilizando o método de variáveis instrumentais não encontramos um coeficiente significante, a princípio isso indicaria uma ausência de relação entre *downloads* e shows. Entretanto, quando olhamos para o valor-p do teste de sobreidentificação das colunas (6) e (7) observamos valores que indicam a rejeição da hipótese nula de sobreidentificação. Isso talvez reflita uma possível correlação dos instrumentos com o gosto por música dos indivíduos e gera uma dúvida sobre as estimativas geradas com uso de tais instrumentos. Como discutimos anteriormente, para explicar o consumo de CDs, como é esperado que o viés seja de sobreestimar o coeficiente, o viés gerado permite ter uma idéia do sinal do efeito. Entretanto, para a análise de bens subordinados e consumo de shows, espera-se que os *downloads* tenham um efeito positivo. Nesse caso, se houver uma ausência de causalidade entre *downloads*, shows e bens subordinados o viés devido aos instrumentos tenderia a gerar uma relação positiva e espúria.

Tabela B.2 Resultados de Primeiro Estágio - Amostra Coletada em Sala de Aula

MÉTODO	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	OLS
VARIÁVEIS	Baixou CD	Baixou CD	Frequência de Download	Faz download pelo menos 1x/semana
Nº de programas p2p	0.0383** (0.0179)	0.0400** (0.0175)	0.307** (0.124)	0.0608*** (0.0161)
Taxa de sucesso em download	0.103*** (0.0136)	0.103*** (0.0135)	0.214*** (0.0570)	0.0495*** (0.0121)
Sabe a velocidade de conexão	0.0272 (0.0401)		0.441** (0.204)	0.0469 (0.0386)
Tempo de acesso à internet em casa	-0.00587 (0.00928)		0.104** (0.0435)	0.0107 (0.00886)
Estatística F (instrumentos excluídos) ¹	17.98	36.02	11.71	10.12
Observações	502	502	502	502
R ²	0.220	0.218	0.150	0.236

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Teste de significância conjunta dos instrumentos utilizados

Na tabela acima vemos que, dentre os instrumentos utilizados nas equações de primeiro estágio, o número de programas p2p e a taxa de sucesso em *download* são as variáveis mais efetivas em explicar as medidas de *download*. Os coeficientes estimados mostram sinais positivos indicando um efeito negativo sobre o custo subjetivo de *download* ou um efeito positivo sobre a habilidade de *download*. O tempo de acesso a Internet e a *dummy* “sabe a velocidade de conexão” não se mostram significantes em nenhuma das especificações em que são utilizadas como instrumentos. A especificação da coluna (2) exclui essas duas variáveis e como consequência o que observamos é um aumento da estatística F de significância global dos instrumentos. Pela regra de bolso de Staiger e Stock (1997) em nenhuma das especificações os instrumentos poderiam ser considerados fracos. Entretanto, esse aumento substancial, de 17,98 para 36,02, na estatística F pode ser um indício de que essas duas variáveis sejam instrumentos fracos.

Os resultados das tabelas acima confirmam a evidência de que os *downloads* possuem um efeito negativo sobre o consumo de CDs. Entretanto, devido ao fato do gosto por música possivelmente influenciar a escolha dos indivíduos com relação às variáveis instrumentais utilizadas, não é possível ter uma certeza acerca dos resultados estimados para bens subordinados e shows.

C - Tabelas Complementares a Abordagem Empírica

Tabela C.1 - Estimativas de Primeiro Estágio - Outras Medidas de *Download* – Mínimos Quadrados Ordinários

	Baixou CD de Música pela Internet			Baixou Música pela Internet			Faz Download de Música Frequentemente		
Variáveis Instrumentais	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Desfragmentou o computador	0.0592*** (0.0134)	0.0665*** (0.0131)	0.0660*** (0.0131)	0.0228 (0.0143)	0.0202 (0.0143)	0.0222 (0.0143)	0.0334** (0.0133)	0.0314** (0.0133)	0.0329** (0.0133)
Formatou o computador	0.0409*** (0.0122)	0.0499*** (0.0115)	0.0475*** (0.0115)	0.0423*** (0.0130)	0.0404*** (0.0130)	0.0396*** (0.0130)	0.0364*** (0.0120)	0.0351*** (0.0120)	0.0339*** (0.0120)
Utiliza Antivírus	0.0469** (0.0214)	0.0571*** (0.0207)	0.0556*** (0.0206)	0.0277 (0.0213)	0.0266 (0.0213)	0.0260 (0.0213)	0.0267 (0.0206)	0.0254 (0.0206)	0.0253 (0.0206)
Conhece o Disco Rígido (HD)	0.0726*** (0.0224)	0.0742*** (0.0224)	0.0729*** (0.0222)	0.0476** (0.0238)	0.0478** (0.0238)	0.0461* (0.0237)	0.0564** (0.0226)	0.0550** (0.0225)	0.0550** (0.0224)
Conhece a Memória Ram	0.0137 (0.0209)	0.0165 (0.0208)	0.0177 (0.0207)	0.00885 (0.0223)	0.00950 (0.0223)	0.0108 (0.0223)	0.0328 (0.0210)	0.0352* (0.0209)	0.0346* (0.0209)
Conhece o Processador	-0.000989 (0.0182)	0.00292 (0.0181)	0.00325 (0.0180)	0.0253 (0.0199)	0.0267 (0.0199)	0.0257 (0.0199)	0.00923 (0.0187)	0.0122 (0.0186)	0.00973 (0.0187)
É contra a Pirataria			-0.0491*** (0.0104)		-0.0451*** (0.0117)			-0.0498*** (0.0108)	
Nº de Produtos de Camêlo ou Feiras			0.0126*** (0.00300)		0.0154*** (0.00350)	0.0187*** (0.00340)		0.0130*** (0.00316)	0.0167*** (0.00308)
Constant	1.039*** (0.0442)	1.045*** (0.0438)	1.015*** (0.0441)	0.915*** (0.0475)	0.894*** (0.0480)	0.880*** (0.0479)	1.045*** (0.0453)	1.030*** (0.0457)	1.013*** (0.0456)
Observações	7003	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.148	0.147	0.154	0.141	0.147	0.145	0.179	0.186	0.183

Desvios Padrões robustos entre Parênteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Teste F de significância conjunta dos instrumentos utilizados

Tabela C.2 - Variáveis Instrumentais (GMM) - Outras Medidas de *Download*

Variáveis Explicativas	Comprou CD			Foi a Show		
	(1)	(2)	(3)	(4)	(5)	(6)
Baixou CD	-0.245*** (0.0816)			0.213*** (0.0727)		
Baixou CD#	-0.237*** (0.0804)			0.245*** (0.0746)		
Baixou Música		-0.357*** (0.117)			0.329*** (0.109)	
Baixou Música#		-0.354*** (0.116)			0.351*** (0.111)	
Baixa Músicas Frequentemente			-0.340*** (0.104)			0.254*** (0.0938)
Baixa Músicas Frequentemente #			-0.329*** (0.104)			0.306*** (0.0981)
Comprou CD pirata	-0.449*** (0.0772)	-0.389*** (0.0857)	-0.406*** (0.0806)	0.0836 (0.0697)	0.0257 (0.0798)	0.0623 (0.0741)
Comprou CD pirata#	-0.381*** (0.0905)	-0.309*** (0.0980)	-0.343*** (0.0929)	0.254*** (0.0861)	0.188** (0.0953)	0.220** (0.0893)
Teste de sobreidentificação (p-valor) ²	0.4479	0.5021	0.6419	0.0052	0.00147	0.0052
Teste de sobreidentificação (p-valor) ² #	0.5970	0.7279	0.7674	0.3223	0.3587	0.2468
Observações	7003	7003	7003	7003	7003	7003
R ²	-0.141	-0.158	-0.130	0.091	0.019	0.070

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por músico

¹ Faz download de músicas pelo menos 1 vez na semana²H0: restrições de sobreidentificação são válidas

: Estimativas excluem do primeiro estágio a dummy: "É contra pirataria"

Tabela C.3 - Efeito de *Downloads* de Música – Outras Medidas - sobre a Probabilidade de Comprar CDs Piratas

Variáveis	Comprou CD Pirata		
	(1)	(2)	(3)
Baixou CD	-0.266*** (0.067)		
Baixou Música		-0.347*** (0.102)	
Faz download Frequentemente ¹			-0.318*** (0.091)
Teste de sobreidentificação (p-valor) ²	0.5983	0.3310	0.2056
Observações	7003	7003	7003
R ²	-0.086	-0.195	-0.127

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal,

Anos de Educação na USP e interesse declarado por música

¹ Faz download de músicas pelo menos 1 vez na semana²H0: restrições de sobreidentificação são válidas

Tabela C.4 - Probabilidade de Comprar Bens Subordinados - GMM

Variáveis	Comprou Bem Subordinado			
	(1)	(2)	(3)	(4)
Utiliza a Internet para baixar Música	-0.00681 (0.0886)			
Baixou CD		-0.00382 (0.0604)		
Baixou Música			-0.0166 (0.0861)	
Faz Download Frequentemente ¹				-0.000603 (0.0761)
Comprou CD pirata	0.137* (0.0730)	0.137* (0.0731)	0.141* (0.0770)	0.136* (0.0744)
Constante	0.310*** (0.113)	0.307*** (0.0807)	0.318*** (0.0919)	0.303*** (0.0933)
Teste de sobreidentificação (p-valor) ²	0.8342	0.8340	0.8391	0.8333
Observações	7003	7003	7003	7003
R ²	0.042	0.042	0.039	0.042

Desvios Padrões robustos entre Parenteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Faz download de músicas pelo menos 1 vez na semana²H0: restrições de sobreidentificação são válidas**Tabela C.5 - Efeito dos Downloads sobre Consumo de CDs e Shows - Análise para Variáveis Abertas(contínuas)**

MÉTODO VARIÁVEIS	(1)	(2)	(3)	(4)
	GMM CDs	GMM CDs	GMM Shows	GMM Shows
Musicas baixadas		-0.0207 (0.0170)		0.0456*** (0.0155)
CDs baixados	-0.0275 (0.0211)		0.0506*** (0.0155)	0.1073 (0.0873)
CDs piratas	-0.345** (0.173)	-0.3518** (0.1653)	0.0900 (0.0916)	
Observações	7003	7003	7003	7003
R ²	-0.383	-0.284	-0.677	-0.3134

Desvios Padrões Robustos entre paratereses

*** p<0.01, ** p<0.05, * p<0.1

Variáveis controles: idade, sexo, cor, renda, instituto,

Anos de Ingresso na Universidade e interesse declarado por música

Tabela C.6 - Frequências Relativas de Consumo por Estilo de Música

Estilos Musicais	Comprou	
	CD	Foi a Show
	%	
jazz	42,26	73,68
soul	39,91	73,24
classica	38,91	64,24
blues	38,85	73,50
mpb	36,93	68,81
alternativa	35,43	71,62
rock	33,80	65,10
samba	33,38	76,39
pop	33,31	61,51
trilha	33,03	62,72
metal	32,19	65,91
romantica	32,01	59,70
reggae	31,86	71,96
rap	28,96	63,58
forro	28,83	72,07
eletronica	28,19	63,81
brega	27,08	70,00
axe	25,96	70,59
pagode	25,00	69,34
sertanejo	24,83	67,41
funk	23,53	73,86

Tabela C.7 - Tabela 6 - Com Controles

Método	MQO	MQO	GMM	GMM	GMM	GMM	GMM	GMM
	Comprou CD					Foi a Show		
Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Utiliza a Internet para Baixar Música	-0.0994***	-0.141***	-0.256**	-0.370***	-0.338***	0.352***	0.293***	0.396***
Comprou CD Pirata	0.0963***	0.107***		-0.426***	-0.374***		0.0768	0.242***
idade em anos	0.00605***	0.00671***	0.00488**	0.00320	0.00370*	-0.000521	-0.00156	0.000127
sexo (masculino)	-0.0512***	-0.0646***	-0.0618***	-0.0666***	-0.0667***	-0.0629***	-0.0583***	-0.0591***
Anos de Estudo na USP	-0.00149	-0.000776	-0.00138	-0.00394*	-0.00365*	0.00545***	0.00581***	0.00693***
Amarelo	-0.0579*	-0.0297	-0.0281	-0.0374	-0.0362	-0.173***	-0.169***	-0.168***
Branco	-0.0225	-0.0152	-0.0133	-0.0161	-0.0161	0.00271	0.00697	0.00456
Indígena	0.125	0.150	0.195*	0.299***	0.285**	0.0152	0.00775	-0.0447
Negro	0.0322	0.0432	0.0504	0.0621	0.0597	0.0293	0.0309	0.0231
Pardo	-0.0251	-0.0152	-0.0153	-0.0174	-0.0171	0.0147	0.0163	0.0159
renda <900 r\$	-0.226***	-0.216***	-0.212***	-0.189***	-0.192***	-0.131***	-0.135***	-0.143***
900 r\$ < renda < 1000 r\$	-0.173***	-0.163***	-0.157***	-0.130**	-0.134**	-0.144***	-0.151***	-0.157***
1000 r\$ < renda < 1200 r\$	-0.221***	-0.214***	-0.203***	-0.160***	-0.166***	-0.0749**	-0.0825**	-0.0981***
1200 r\$ < renda < 1400 r\$	-0.196***	-0.187***	-0.182***	-0.163***	-0.166***	-0.112***	-0.116***	-0.122***
1400 r\$ < renda < 1800 r\$	-0.217***	-0.200***	-0.196***	-0.187***	-0.190***	-0.147***	-0.147***	-0.150***
1800 r\$ < renda < 2200 r\$	-0.152***	-0.143***	-0.140***	-0.135***	-0.136***	-0.138***	-0.137***	-0.139***
2200 r\$ < renda < 3000 r\$	-0.126***	-0.117***	-0.111***	-0.0965***	-0.0999***	-0.109***	-0.109***	-0.116***
3000 r\$ < renda < 4000 r\$	-0.101***	-0.0911***	-0.0857***	-0.0704**	-0.0730**	-0.114***	-0.114***	-0.122***
4000 r\$ < renda < 5000 r\$	-0.0991***	-0.0883***	-0.0798***	-0.0585*	-0.0620**	-0.100***	-0.100***	-0.112***
5000 r\$ < renda < 7000 r\$	-0.0887***	-0.0782***	-0.0710**	-0.0520*	-0.0555*	-0.0684**	-0.0687**	-0.0777***
7000 r\$ < renda < 10000 r\$	-0.0414	-0.0393	-0.0329	-0.0217	-0.0246	-0.0631**	-0.0620**	-0.0678**
10000 r\$ < renda < 15000 r\$	-0.0450	-0.0375	-0.0377	-0.0336	-0.0342	-0.0336	-0.0363	-0.0359
EACH	0.0821***	0.0769***	0.0832***	0.105***	0.103***	-0.0618**	-0.0645**	-0.0738***
ECA	0.129***	0.0985***	0.0922***	0.0628	0.0664*	0.0172	0.0260	0.0346
EE	-0.0247	-0.0149	-0.00968	0.0261	0.0217	-0.101**	-0.113***	-0.123***
EEFE	-0.0491	-0.0292	-0.0241	-0.00550	-0.00534	-0.0696	-0.0679	-0.0817
FAU	0.0510	0.0453	0.0385	0.00390	0.00758	-0.00559	0.00326	0.0147
FCF	0.0290	0.0324	0.0272	0.0143	0.0161	0.00184	0.00700	0.00925
FD	0.176***	0.174***	0.173***	0.169***	0.170***	-0.0496	-0.0497	-0.0475
FE	0.221***	0.193***	0.205***	0.220***	0.218***	-0.0234	-0.0222	-0.0302
FEA	0.00602	0.0137	0.0155	0.0247	0.0235	-0.0369	-0.0402*	-0.0422*
FFLCH	0.134***	0.108***	0.105***	0.0854***	0.0876***	0.0211	0.0281	0.0323
FM	0.0923***	0.102***	0.103***	0.103***	0.104***	-0.0178	-0.0173	-0.0188
FMVZ	0.158***	0.169***	0.159***	0.120*	0.126**	0.0115	0.0175	0.0324
FO	-0.0186	0.00146	0.00314	0.00280	0.00387	-0.129***	-0.126***	-0.129***
FSP	0.0319	0.0492	0.0500	0.0293	0.0312	-0.0829*	-0.0741*	-0.0709
FZEA	-0.0295	-0.0253	-0.00657	0.0456	0.0384	-0.0758	-0.0831*	-0.105**
IAG	0.133**	0.133**	0.135**	0.136*	0.135*	-0.0668	-0.0626	-0.0669
IB	0.0307	0.00867	0.00688	-0.00473	-0.00308	0.0426	0.0480	0.0493
ICB	0.0108	0.00298	0.0173	0.0616	0.0552	-0.0272	-0.0289	-0.0508
ICMC	-0.00376	0.0105	0.000265	-0.0304	-0.0278	-0.123***	-0.117***	-0.104***
IF	0.0636*	0.0543*	0.0466	0.0213	0.0252	-0.0998***	-0.0968***	-0.0865**
IGC	0.0585	0.0604	0.0557	0.0444	0.0479	0.0197	0.0234	0.0246
IME	0.0887**	0.0947***	0.0914**	0.0726*	0.0753*	-0.0906**	-0.0874**	-0.0797**
IO	-0.0444	-0.0433	-0.0455	-0.0486	-0.0488	0.0594	0.0659	0.0625
IP	0.0264	0.0182	0.00891	-0.0239	-0.0195	-0.0617	-0.0554	-0.0421
IQ	0.0636	0.0605	0.0623	0.0618	0.0621	-0.179***	-0.178***	-0.180***
IRI	0.0574	0.0476	0.0556	0.0838	0.0806	-0.0588	-0.0538	-0.0743
Poli	0.0320	0.0364	0.0313	0.00959	0.0127	-0.116***	-0.112***	-0.104***
Não Interessado em Música		-0.381***	-0.419***	-0.440***	-0.430***	-0.321***	-0.350***	-0.317***
Pouco Interessado em Música		-0.239***	-0.251***	-0.245***	-0.243***	-0.226***	-0.239***	-0.233***
Interessado em Música		-0.131***	-0.133***	-0.125***	-0.125***	-0.0971***	-0.102***	-0.103***
Constante	0.373***	0.500***	0.660***	0.866***	0.818***	0.587***	0.649***	0.494***
Observações	7003	7003	7003	7003	7003	7003	7003	7003
R ²	0.063	0.095	0.082	-0.099	-0.062	0.054	0.068	0.021

Desvios Padrões robustos entre Parênteses

*** p<0.01, ** p<0.05, * p<0.1

Controles: idade, sexo, cor, instituto (USP), renda familiar mensal, Anos de Educação na USP e interesse declarado por música

¹ Faz download de músicas pelo menos 1 vez na semana²H0: restrições de sobreidentificação são válidas

D - Questionário

As respostas inseridas neste questionário serão utilizadas apenas para fins de pesquisa acadêmica, sendo vedado seu uso para qualquer outro fim.

*** Required**

Qual é a sua idade em anos ? ***Digite em número ex : 15**

1. Informe o seu Sexo: ***Escolha uma das duas opções abaixo**

- ☐ Masculino
- ☐ Feminino

2. Informe sua Cor ou Raça (Predominante) ***Escolha uma das 5 opções abaixo.**

- ☐ Branco
- ☐ Negro
- ☐ Amarelo
- ☐ Pardo
- ☐ Indígena
- ☐ Outros
- ☐ Prefiro não declarar

3. Você reside com seus pais ou responsáveis? ***Escolha uma das duas opções abaixo**

- ☐ Sim
- ☐ Não

4. Aproximadamente, qual é a sua renda familiar mensal em reais? ***Escolha uma das opções abaixo**

- ☐ Menos de 900
- ☐ de 900 - 1000
- ☐ de 1000 - 1200
- ☐ de 1200 - 1400
- ☐ de 1400 - 1800
- ☐ de 1800 - 2200
- ☐ de 2200 - 3000
- ☐ de 3000 - 4000
- ☐ de 4000 - 5000
- ☐ de 5000 - 7000
- ☐ de 7000 - 10000
- ☐ de 10000 - 15000

- ☐ acima de 15000

5. Qual é o seu vínculo com a USP? *Escolha uma das opções abaixo

- ☐ Estudante - Graduação
- ☐ Estudante - Pós-graduação
- ☐ Docente / Pesquisador
- ☐ Funcionário
- ☐ Outro

6. Informe o instituto onde você estuda ou trabalha.

7. Qual foi o ano do seu ingresso na USP? *Informe o ano. Ex: 2008

8. Aproximadamente, quantos produtos com o nome de um artista musical você comprou no último mês? (EX: Camisetas, bonés, adesivos e etc) (Exclua Shows, CDs e DVDs.) *Responda em número. EX: 13

9. Aproximadamente, quantos produtos com o nome de um artista musical você comprou nos últimos 6 meses? (EX: Camisetas, bonés, adesivos e etc) (Exclua Shows, CDs e DVDs.) Responda em número. EX: 12

10. Aproximadamente, quantos reais você gastou nos últimos 6 MESES com os produtos referentes à questão anterior? Responda em números: Ex: 350

11. Aproximadamente, quantos CDs piratas você comprou no último mês (adquiridos numa banca de camelô ou de um vendedor ambulante)? Responda em números: Ex: 350

12. Aproximadamente, quantos CDs piratas você comprou nos últimos 6 MESES (adquiridos numa banca de camelô ou de um vendedor ambulante)? Responda em números: Ex: 350

13. Aproximadamente, quantos CDs originais de música você comprou no último mês? Responda em números: Ex: 350

14. Aproximadamente, quantos CDs originais de música você comprou nos últimos 6 MESES? Responda em números: Ex: 350

15. Você possui um computador em casa? (Caso sua resposta seja "Não", pule a próxima pergunta) *Escolha uma das opções abaixo.

- ☐ Sim
- ☐ Não

16. Com quantas pessoas você compartilha o seu computador em casa? *Responda em número. Ex: 5

17. Você possui um computador portátil (laptop ou notebook) com conexão wireless ?Escolha uma das opções abaixo

- ☐ Sim
- ☐ Não

18. Você possui um tocador de MP3 portátil (MP3 player) ?Escolha uma das opções abaixo

- ☐ Sim
- ☐ Não

19. Você possui conexão à Internet em casa? *Escolha uma das opções

- ☐ Sim
- ☐ Não

20. Qual é o tipo da sua conexão à Internet?Escolha uma das opções abaixo

- ☐ Discada
- ☐ Banda Larga
- ☐ Wi -Fi
- ☐ Satélite
- ☐ 3G - Celular
- ☐ Outros
- ☐ Não Sei
- ☐ Não tenho acesso à Internet em casa

21. Informe a velocidade de sua conexão.Escolha uma das opções abaixo

- ☐ menor que 1 mega
- ☐ 1 mega - 3 mega
- ☐ 3 mega - 6 mega
- ☐ 6 mega - 8mega
- ☐ acima de 8 mega
- ☐ Não sei
- ☐ Não tenho acesso à Internet em casa

22. Há quanto tempo você possui acesso à Internet em casa?Escolha uma das opções abaixo.

- ☐ 0 - 6 meses
- ☐ 6 - 12 meses
- ☐ 1 - 2 anos

- ☐ 2 - 3 anos
- ☐ 3 - 5 anos
- ☐ 5 - 8 anos
- ☐ 8 - 10 anos
- ☐ 12 - 15 anos
- ☐ + de 15 anos
- ☐ Não tenho acesso à Internet em casa

23. Você possui acesso à Internet em outros locais como Universidade ou trabalho? Escolha uma das opções abaixo

- ☐ Sim
- ☐ Não

24. Você utiliza a Internet para baixar músicas (*download*)? Escolha uma das opções abaixo

- ☐ Sim
- ☐ Não

25. Você lê ou já leu alguma revista especializada em informática? *Escolha uma das opções abaixo

- ☐ Sim
- ☐ Não

26. Se você tivesse que se posicionar politicamente, você diria que na maior parte dos assuntos você é: Escolha uma das opções abaixo

- ☐ De Direita
- ☐ De Centro
- ☐ De Esquerda
- ☐ Outros
- ☐ Não Sei

27. Se você tivesse que se posicionar politicamente, você diria que na maior parte dos assuntos você tem mais afinidade com a posição do seguinte partido: Escolha uma das opções abaixo

- ☐ PDT
- ☐ PC do B
- ☐ PR
- ☐ DEM
- ☐ PMDB

- ☐ PPS
- ☐ PP
- ☐ PSDB
- ☐ PSB
- ☐ PT
- ☐ PSTU
- ☐ PV
- ☐ PTB
- ☐ PCB
- ☐ PSOL
- ☐ PRTB
- ☐ PTdoB
- ☐ PTN
- ☐ PTC
- ☐ PSL
- ☐ PSC
- ☐ PSDC
- ☐ PMN
- ☐ outros
- ☐ Nenhum

28. Você é a favor do aumento de restrições sobre o porte de armas no Brasil? Escolha uma das opções

- ☐ Sim
- ☐ Não
- ☐ Indiferente
- ☐ Não sei

29. Você é a favor da legalização da Maconha no Brasil? Escolha uma das opções

- ☐ Sim
- ☐ Não
- ☐ Indiferente
- ☐ Não sei

30. Você é a favor da maior severidade da Legislação de trânsito no tocante a direção alcoolizada? Escolha uma das opções

- ☐ Sim
- ☐ Não
- ☐ Indiferente
- ☐ Não sei

31. O que você diria do crime organizado e o comércio de produtos falsificados? Escolha uma das opções abaixo

- ☐ São relacionados
- ☐ São não relacionados
- ☐ Não sei dizer

32. Qual é a sua posição com respeito ao comércio de produtos pirateados? Escolha uma das opções

- ☐ A favor
- ☐ Contra
- ☐ Indiferente
- ☐ Não tenho posição definida

33. Quais dos seguintes produtos, da lista abaixo, você já comprou numa banca de camelô?

- ☐ Camisas Esportivas
- ☐ Bolsas
- ☐ Mochilas
- ☐ Tênis
- ☐ Roupas
- ☐ Celular
- ☐ Acessórios para aparelhos eletrônicos
- ☐ Produtos Esportivos
- ☐ DVD
- ☐ Relógio
- ☐ Óculos
- ☐ Sapato
- ☐ Jóias (Bijuterias)
- ☐ Softwares de Computador
- ☐ Jogos de Azar (Baralho, Dominó)
- ☐ Jogos de Videogame
- ☐ CDs de música
- ☐ Outros

34. Aproximadamente, quantos CDs de música você baixou pela Internet no último mês? *Responda em números. Ex: 4

35. Aproximadamente, quantos CDs de música você baixou pela Internet nos últimos 6 MESES ? *Responda em números. Ex: 4

36. Aproximadamente, quantas músicas você baixou pela Internet nos últimos 15 dias ? *Responda em números. Ex: 4

37. Aproximadamente, a quantos Shows de música você foi no último mês? *Responda em números. Ex: 4

38. Aproximadamente, a quantos Shows de música você foi nos últimos 6 MESES? *Responda em números. Ex: 4

39. Aproximadamente, quanto você gastou, em reais, com os shows de música nos últimos 6 meses? Responda em número. Ex: 120

40. Selecione abaixo os programas e sites abaixo que você utiliza ou já utilizou para baixar música pela Internet? Caso você nunca tenha feito *download*, prossiga para próxima questão

- ☐ Lime Wire
- ☐ Kazaa
- ☐ Itunes
- ☐ Warez
- ☐ Ares
- ☐ Napster
- ☐ Azureus
- ☐ mininova.org
- ☐ emule
- ☐ orkut.com
- ☐ u-torrent
- ☐ piratebay.org
- ☐ 4shared
- ☐ Outros

41. Quais das seguintes tarefas você já realizou quando utilizou um computador?

- ☐ Desfragmentador de Discos
- ☐ Formatar o Computador
- ☐ Trocar componentes de Hardware

- ☐ Instalar Software
- ☐ Rodar Antivírus
- ☐ Nenhuma das Opções

46. Quais dos seguintes itens abaixo, referentes à configuração do seu computador, você conhece?

- ☐ Memória Ram
- ☐ Memória em HD
- ☐ Processador
- ☐ Não conheço a configuração do meu computador
- ☐ Não tenho computador

47. Quantas vezes por semana você baixa músicas na Internet? Escolha uma das opções abaixo

- ☐ Nenhuma
- ☐ de 1 - 2
- ☐ de 2 - 3
- ☐ de 3 - 4
- ☐ de 4 - 5
- ☐ de 5 - 6
- ☐ de 6 - 7
- ☐ de 7 - 8
- ☐ de 8 - 9
- ☐ de 9 - 10
- ☐ + de 10

48. Selecione os DIAS DA SEMANA em que você baixa músicas pela Internet.

- ☐ Nenhum
- ☐ de 2° a 6°
- ☐ Sábados
- ☐ Domingos

50. Em geral, quando você inicia o *download* de um arquivo você consegue concluir o *download*? Escolha uma das opções abaixo

- ☐ Nunca

- ☐ de 0 - 25% das vezes
- ☐ de 25% - 50 % das vezes
- ☐ de 50% - 75% das vezes
- ☐ de 75% - 100% das vezes
- ☐ Sempre
- ☐ Não sei
- ☐ Não utilizo à Internet para fazer *downloads*

51. Quais dos seguintes estilos musicais você geralmente escuta e (ou) baixa na Internet?

- ☐ Rock
- ☐ Reggae
- ☐ Hip- Hop
- ☐ Romântica
- ☐ Pagode
- ☐ Axé
- ☐ Samba
- ☐ Funk-Carioca
- ☐ Sertanejo
- ☐ Clássica
- ☐ Pop
- ☐ Jazz
- ☐ Soul
- ☐ Brega
- ☐ Trilha Sonora
- ☐ Metal
- ☐ Alternativa
- ☐ Eletrônica
- ☐ Tema de Novela
- ☐ Forró
- ☐ Blues
- ☐ MPB
- ☐ Outros

52. Você toca ou já tocou algum instrumento musical ?Escolha uma das opções

- ☐ Sim

- ☐ Não

53. Você participa ou já participou de alguma banda ou projeto musical? Escolha uma das opções

- ☐ Sim
- ☐ Não

54. Como você classifica o seu interesse por música? Escolha uma das opções abaixo

- ☐ Não interessado
- ☐ Pouco Interessado
- ☐ Interessado
- ☐ Muito Interessado