

FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE PÓS GRADUAÇÃO EM ECONOMIA
MESTRADO EM FINANÇAS E ECONOMIA EMPRESARIAL

PEDRO RICARDO BERNARDINO DE FREITAS

**O PAPEL DA ELASTICIDADE DA RENDA TRIBUTÁVEL NA AVALIAÇÃO DO
CUSTO DE EFICIÊNCIA DA TRIBUTAÇÃO**

Rio de Janeiro

2011

PEDRO RICARDO BERNARDINO DE FREITAS

**O PAPEL DA ELASTICIDADE DA RENDA TRIBUTÁVEL NA AVALIAÇÃO DO
CUSTO DE EFICIÊNCIA DA TRIBUTAÇÃO**

Dissertação submetida ao corpo docente da Escola de Pós-Graduação em Economia da Fundação Getúlio Vargas – EPGE / FGV como parte dos requisitos necessários à obtenção do grau de Mestre em Finanças e Economia Empresarial, sob a orientação do Professor Carlos Eugênio Ellery Lustosa da Costa.

Rio de Janeiro

2011

Freitas, Pedro Ricardo Bernardino de

**O papel da elasticidade da renda tributável na avaliação do custo de
eficiência da tributação / Pedro Ricardo Bernardino de Freitas. - 2011.
89 f.**

**Dissertação (mestrado) - Fundação Getulio Vargas, Escola de Pós-
Graduação em Economia.**

Orientador: Carlos Eugênio Ellery Lustosa da Costa.

Inclui bibliografia.

**1. Imposto de renda. 2. Elasticidade-renda. 3. Impostos. 4. Política
tributária. I. Costa, Carlos Eugênio da. II. Fundação Getulio Vargas.
Escola de Pós-Graduação em Economia. III. Título.**

CDD – 336.2

PEDRO RICARDO BERNARDINO DE FREITAS

**O PAPEL DA ELASTICIDADE DA RENDA TRIBUTÁVEL NA AVALIAÇÃO DO
CUSTO DE EFICIÊNCIA DA TRIBUTAÇÃO**

Dissertação submetida ao corpo docente da
Escola de Pós-Graduação em Economia da
Fundação Getúlio Vargas – EPGE / FGV como
parte dos requisitos necessários à obtenção
do grau de Mestre em Finanças e Economia
Empresarial.

Aprovada em 29 de agosto de 2011

BANCA EXAMINADORA

Prof. Carlos Eugênio Ellery Lustosa da Costa -Orientador
(EPGE/FGV)

Prof. Afonso Arinos de Mello Franco Neto
(EPGE/FGV)

Prof. Paulo César Coimbra Lisboa
(UFJF/MG)

À minha querida esposa Jaíse, grande motivadora desta empreitada. Sem seu amor e seu apoio este trabalho não teria um início e muito menos um final.

Ao meu filho João Pedro, que hoje, ao desenhar suas primeiras letras, inicia sua jornada pelos caminhos do conhecimento.

AGRADECIMENTOS

Agradeço aos meus pais e a minha irmã, que ao longo da vida me ensinaram o amor pelos livros, e o valor da educação, da cultura e do estudo;

Ao Professor Carlos Eugênio, que aceitou orientar meus estudos mesmo estando à época no exterior, agradeço os valiosos ensinamentos que recebi, sempre transmitidos de forma serena e cordial;

Ao corpo funcional da Escola de Pós Graduação em Economia da Fundação Getúlio Vargas, e particularmente ao Vitor Barros Souza, (o “anjo da guarda” da EPGE!), pelo suporte e apoio recebido em todos os momentos ao longo do curso;

Aos colegas da Receita Federal do Brasil pelo apoio recebido em momentos cruciais desta caminhada, especialmente Adriana Moraes Marques Sanan, Célio Ramos da Silva, José Antônio Barbosa, Zenilda Dal’Col Benevides, Laura Gadelha Xavier, Gelson Machado Guarçoni e Silvio Magno da Silva Alves.

“...Vivendo se aprende, mas o que se aprende mais é só a fazer outras maiores perguntas...”

João Guimarães Rosa

RESUMO

O trabalho utiliza conceitos da economia do bem estar e a elasticidade da renda tributável para analisar o custo social das reações comportamentais dos contribuintes do imposto de renda sobre as pessoas físicas no Brasil, em resposta a uma mudança de política tributária. A elasticidade da renda tributável despertou grande atenção recente, motivada pela perspectiva de estimar em conjunto todas as reações comportamentais a uma mudança no sistema tributário através de um único parâmetro, mensurando custos de eficiência e de bem estar suportados pela economia de forma relativamente simples. O trabalho aborda a utilização de medidas de variação de bem estar para avaliar mudanças na política tributária e faz uma resenha da literatura sobre a elasticidade da renda tributável, conceitos, características, vantagens e limitações. Um modelo de preferências é especificado para exemplificar a dimensão das reações à tributação e os custos de eficiência envolvidos, e discutir a viabilidade do emprego da elasticidade da renda tributável como parâmetro estrutural.

Palavras-chave: elasticidade da renda tributável; carga tributária; perda de peso morto; imposto de renda sobre pessoa física; reação comportamental.

ABSTRACT

This study aims to analyse the social cost and behavioural responses due the taxation of earned income in Brazil using welfare concepts and the elasticity of taxable income with respect to the net of tax rate. The elasticity of taxable income has deserved great attention and is a main issue in public economics research agenda. Under some conditions, it holds the perspectives of measuring efficiency costs of income taxation from a wide array of behavioural responses just estimating one single parameter. The study introduces the use of welfare changes measures for tax policy evaluation and reviews the concepts about elasticity of taxable income, its characteristics, advantages and limitations of use. We develop a model in which we show the dimension of responsiveness of taxpayers. Based on this model, we estimate the efficiency cost and excess burden of income taxation and discuss the use of elasticity of taxable income as a structural parameter.

Keywords: elasticity of taxable income; tax burden; deadweight loss; personal income tax; behavioural response.

SUMÁRIO

1 INTRODUÇÃO	11
2 A TRIBUTAÇÃO SOBRE A RENDA DA PESSOA FISICA NO BRASIL – BREVE HISTORICO E CARACTERISTICAS BASICAS	14
2.1 Breve Histórico	14
2.2 Características Básicas	17
2.3 Alíquotas e Progressividade	21
2.4 Participação Relativa do Imposto de Renda – Pessoa Física no Sistema Tributário Nacional	23
3 CUSTOS DE EFICIENCIA E A ELASTICIDADE DA OFERTA DE TRABALHO	25
3.1 Tributação Sobre a Renda e os Custos de Eficiência	25
3.2 Tributação Sobre a Renda e Oferta De Trabalho	30
3.2.1 Modelo Básico	30
3.2.2 Modelo Não Linear	33
3.2.3 Modelo Não Convexo	38
3.2.4 Equação de Slutsky Para a Oferta De Trabalho	40
3.2.5 Estimativas de Elasticidade da Oferta de Trabalho	42
3.2.6 A Elasticidade da Oferta de Trabalho - Limitações Ao Estimar Custos de Eficiência	44
4 A ELASTICIDADE DA RENDA TRIBUTAVEL (ETI)	46
4.1 Modelo Básico	46

4.2 Estudos Empíricos - Resultados Encontrados Na Literatura	50
4.3 Restrições e Limitações	52
4.4 Elisão e Evasão Fiscal	53
4.5 Modelo de Chetty – Generalização do Modelo Básico	55
4.6 Avaliação De Bem Estar Utilizando o Modelo De Chetty	57
5 ESPECIFICAÇÃO DE UM MODELO ILUSTRATIVO DA TRIBUTAÇÃO SOBRE A RENDA NO BRASIL CONSIDERANDO O COMPORTAMENTO DOS INDIVÍDUOS ..	60
5.1 Derivação do modelo básico	60
5.2 Estimativa de uma alíquota marginal única para o IRPF	64
5.3 Efeitos sobre a arrecadação e perda de peso morto	66
5.4 Resultados	69
6 CONCLUSÃO	75
REFERÊNCIAS BIBLIOGRÁFICAS	78
APÊNDICE A	82
APÊNDICE B	84
APÊNDICE C	86
APÊNDICE D	89

1 INTRODUÇÃO

O imposto sobre a renda das pessoas físicas é seguramente um dos tributos mais sujeitos a discussões, questionamentos e a intensa pressão política por parte dos diversos segmentos da sociedade. A tributação sobre a renda dos indivíduos, acompanhada da progressividade das alíquotas do imposto em relação ao nível de renda, é geralmente percebida como uma forma de tributação socialmente justa e um mecanismo de redistribuição de renda e redução de desigualdades. Todavia, o imposto de renda é distorcivo e reduz a eficiência da economia.

Em que pese a desejável busca por uma maior eficiência tributária, o imposto de renda agrega como em nenhum outro tributo as exigências de tratamento equânime quanto à distribuição do encargo relativo ao seu pagamento. A pressão da sociedade é mais contundente e motivadora de boa parte das preocupações ponderadas pelos agentes políticos responsáveis pelo desenho e pela administração do sistema tributário.

O imposto reduz a disponibilidade de renda do indivíduo. Uma decisão governamental que resulte em uma majoração de alíquota provoca mudanças de comportamento no contribuinte, e dependendo da dimensão dos efeitos resultantes do aumento da carga tributária, pode ser fonte de maior ou menor distorção para a economia do país.

Impostos mais altos reduzem a disponibilidade monetária do indivíduo, que pode decidir trabalhar mais para restaurar a renda líquida perdida, o que em si não causa ineficiência econômica. Apesar do ônus representado pelo aumento da carga tributária suportada pelo contribuinte, a redução da renda resulta em um custo privado, e não em um custo social, supondo que o valor das exações recolhidas pelo governo será integralmente revertido à sociedade. Este comportamento decorre do chamado efeito renda resultante da imposição do tributo pelo Estado.

Por outro lado nem todos os bens de consumo e atividades existentes são tributados da mesma forma, o que leva à alteração dos preços relativos entre os bens. Ao invés de

trabalhar mais ou ter menos horas de lazer, o contribuinte do imposto busca fugir da tributação através do consumo preferencial de bens, serviços, e a prática de atividades sujeitas a uma tributação menos gravosa, em detrimento de outros bens, serviços ou atividades expostos a uma carga maior de tributos. Esta conduta é resultado do chamado efeito substituição, que é indicativo das ineficiências e perdas econômicas decorrentes de um aumento da tributação. Quanto mais intensa a resposta do contribuinte, maior a perda resultante da carga tributária. É usual o emprego da equação de Slutsky e de elasticidades ao avaliar a dimensão dos efeitos renda e substituição na economia.

No caso da tributação sobre a renda, a análise dos reflexos resultantes de políticas tributárias sobre o comportamento dos indivíduos e das distorções causadas por estas políticas passou originalmente pelos clássicos conceitos da economia do trabalho, e teve como principal parâmetro de avaliação a chamada elasticidade da oferta de trabalho (*elasticity of labor supply*). Porém evidências empíricas apontam para um valor muito pequeno para a elasticidade compensada da oferta do trabalho; o valor implicaria um comportamento neutro do contribuinte em relação à tributação, o que contraria o senso comum. A partir do trabalho de Feldstein (1995, 1999), ganhou corpo uma nova frente de pesquisa em finanças públicas, tendo por base a avaliação de todas as possíveis reações dos contribuintes à tributação sobre a renda utilizando um único parâmetro, a elasticidade da renda tributável (*elasticity of taxable income*). A estimativa das variações percentuais da renda tributável em relação ao percentual de mudança da alíquota marginal do imposto permitiria produzir avaliações mais consistentes de custos sociais e de bem estar, e consequentemente a própria estratégia governamental adotada em relação à política tributária.

O trabalho está dividido em cinco partes. Na primeira parte, são apresentadas a definição, a evolução histórica e as características básicas do imposto sobre a renda das pessoas físicas no Brasil. A segunda parte traz o conceito do custo de eficiência da tributação sobre a renda e é apresentada a elasticidade da oferta de trabalho. Na terceira parte, são apresentadas as características e os modelos em que se baseia a elasticidade da renda tributável; São comentadas as principais estimativas das

elasticidades presentes na literatura de finanças públicas, bem como os problemas e limitações típicos de cada tipo de elasticidade. Na quarta parte é especificado um modelo ilustrativo da tributação sobre a renda da pessoa física tendo como base o caso brasileiro. A partir do modelo, são apresentados os cálculos de elasticidades, da alíquota marginal única equivalente ao sistema de alíquotas aplicado sobre o contribuinte pessoa física no Brasil, e resultados referentes a custos de eficiência e variação da arrecadação decorrentes de um hipotético aumento do tributo sobre a renda. Na quinta parte é apresentada a conclusão do trabalho.

2 A TRIBUTAÇÃO SOBRE A RENDA DA PESSOA FÍSICA NO BRASIL – BREVE HISTÓRICO E CARACTERÍSTICAS BÁSICAS

2.1 Breve Histórico

Comparado com outras modalidades de tributação, a implantação do imposto sobre a renda é relativamente recente, e motivou vigorosa reação dos potenciais contribuintes nos países onde a idéia originalmente foi posta em prática. No Reino Unido, primeira nação a adotar esta forma de tributação, o imposto de renda foi criado para financiar os altos dispêndios decorrentes das Guerras Napoleônicas. A sua manutenção em tempos de paz somente ocorreu após muita controvérsia e às custas da impopularidade de seus defensores. No Estados Unidos a polêmica foi materializada em um *imbroglio* jurídico; em princípio declarado inconstitucional pela Suprema Corte de Justiça americana, somente em 1913, após uma emenda à constituição autorizando a tributação sobre a renda, foi instituído o imposto de renda naquele país.

No Brasil, o imposto de renda foi instituído através da Lei 4.625 de 31 de dezembro de 1922, que definiu o orçamento da República para 1923. A idéia de criar um tributo incidente sobre a renda e proventos enfrentou desde o início tenaz oposição; foi precedida por décadas de acalorada discussão e por tentativas infrutíferas de criar um imposto sobre a renda das pessoas físicas e jurídicas no país. Os que argumentavam a favor tinham o imposto de renda como uma forma de tributação mais justa e eqüitativa. Aqueles que eram contrários à nova tributação alegavam o excessivo custo administrativo para promover sua arrecadação e elevada complexidade. Defenderam a criação do novo imposto figuras ilustres do pensamento nacional como Rui Barbosa, sem no entanto lograr êxito imediato.

Ao longo da última década, a participação do imposto de renda sobre pessoas físicas (IRPF) junto à arrecadação de receitas federais brasileiras¹ oscilou entre 12 e 14 %, o

¹ O valor total do imposto de renda sobre pessoas físicas arrecadado agrega o imposto retido na fonte sobre rendimentos do trabalho, o imposto pago na declaração de ajuste anual (DIRPF) e o imposto de renda sobre pessoas físicas cobrado mensalmente (carnê leão).

que representa em torno de 6 a 7 % da arrecadação de tributos no Brasil, e corresponde a um valor próximo entre 2 e 2,5% do PIB brasileiro.

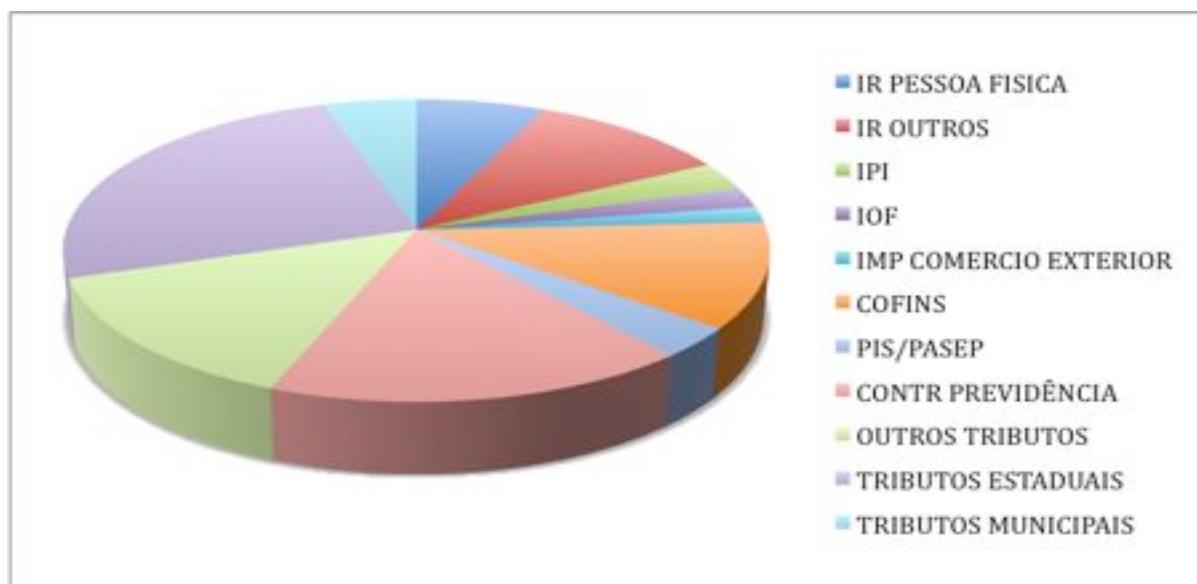


Figura 1 – Participação relativa de cada tributo na na receita tributária brasileira – fonte: Receita Federal do Brasil.

A participação do imposto de renda sobre pessoas físicas face a carga tributária total no Brasil é bem menor do que a registrada na maioria dos países desenvolvidos, onde a participação do imposto frequentemente ultrapassa os 30% da carga tributária, o que acontece, por exemplo, nos Estados Unidos, Canadá, países do norte da Europa, Nova Zelândia e Austrália. Entre os países da OCDE (Organização para a Cooperação e Desenvolvimento Econômico), o percentual de participação do imposto de renda no total da carga tributária oscila em torno de 26%.

O sistema tributário brasileiro ao longo da história não logrou privilegiar a equidade como característica fundamental. O Estado brasileiro há tempos adotou como padrão um modelo de tributação que objetiva um reduzido custo administrativo e político, o que coloca o modelo de tributação sobre a renda dos indivíduos em uma posição pouco vantajosa.

A ordem tributária vigente após a Constituição de 1988 de certa forma agravou a situação. O legislador primário incluiu tributos como o Imposto de Renda (IR) e o

Imposto sobre Produtos Industrializados (IPI) nas disposições que tratam da transferência de receitas tributárias do ente federal para estados e municípios. Como resultado, tributos não sujeitos a esta transferência de recursos, como é o caso das contribuições incidentes sobre o faturamento das empresas, foram privilegiados em detrimento do Imposto de Renda e do IPI.

Nas últimas duas décadas a participação do imposto de renda sobre pessoas físicas em relação à arrecadação total de tributos permaneceu estabilizada; entretanto a participação da Contribuição para o Programa de Integração Social (PIS) e Contribuição para o Financiamento da Seguridade Social (COFINS) mais que dobrou. Outras contribuições federais como a Contribuição de Intervenção no Domínio Econômico (CIDE) e a Contribuição Provisória sobre as Movimentações Financeiras (CPMF) foram criadas e ganharam importância dentro do sistema tributário brasileiro, partindo da mesma filosofia, que engloba a facilidade e baixo custo de arrecadação e a não divisão de receitas tributárias adicionais com outros entes federativos.

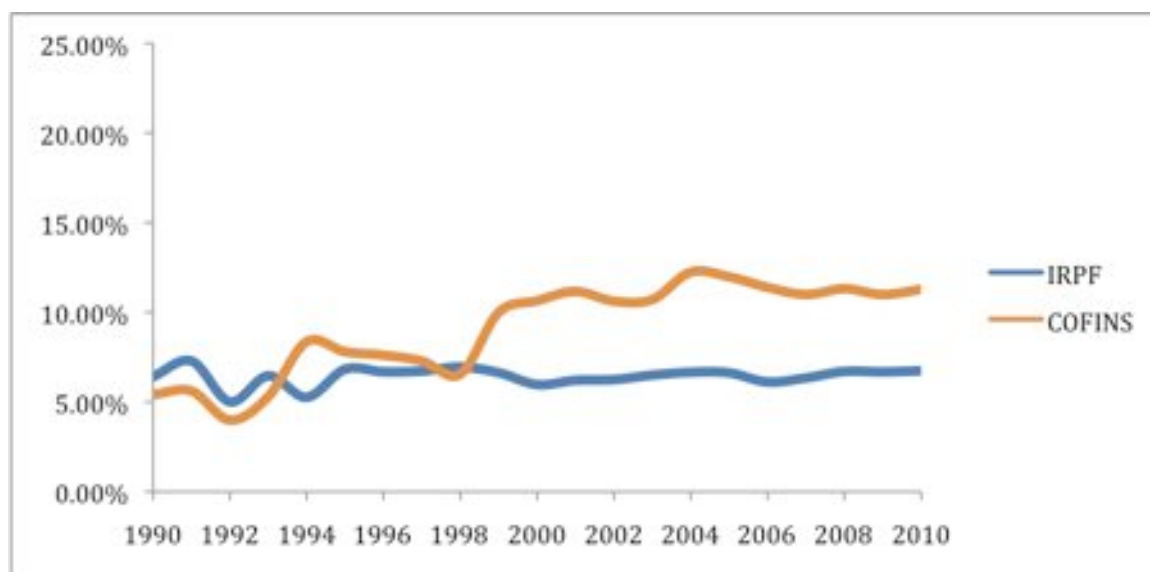


Figura 2: Receita tributária com IRPF e Cofins (em percentual de participação na arrecadação de tributos) – fonte: Receita Federal do Brasil.

Nos últimos quinze anos, a legislação e a regulamentação do imposto de renda pessoa física gozaram de relativa estabilidade, com poucas e pontuais alterações. Se por um lado a manutenção favoreceu uma maior transparência e maior compreensão das

regras básicas por parte do contribuinte, por outro lado, consolidou a baixa participação deste imposto em relação ao valor total de tributos arrecadado anualmente no Brasil.

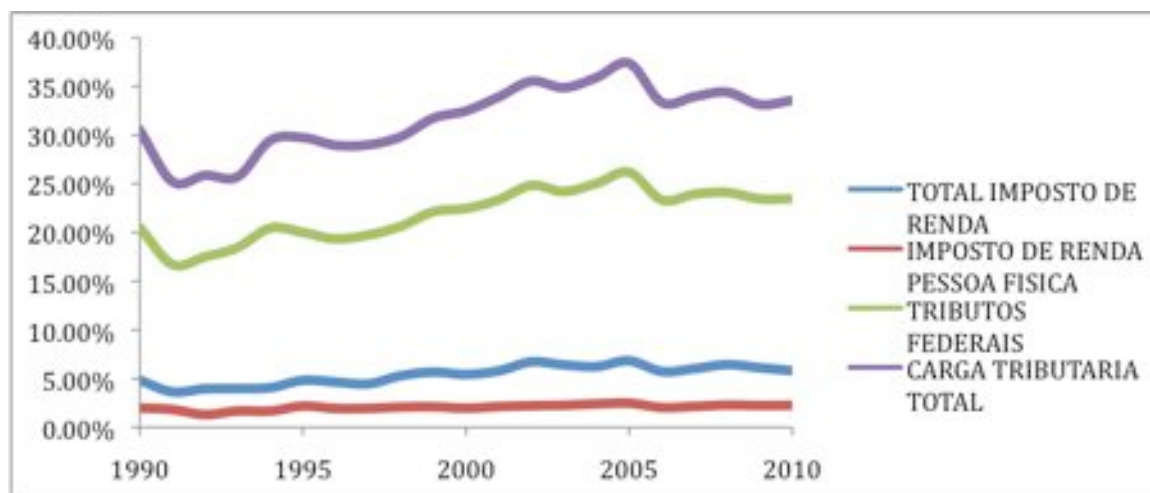


Figura 3 - Receita tributária brasileira em percentual do PIB² – fonte: Receita Federal do Brasil.

2.2 Características Básicas

Observar o conceito de justiça socialmente aceito é uma premissa importante para a construção de instituições sólidas, o que inclui definir uma organização ótima do sistema tributário. Aliomar Baleeiro elencou entre as características desejáveis de um sistema tributário “ser compatível com as idéias de justiça da época”. Na grande maioria das sociedades é tido como justo e razoável que aqueles que gozam de capacidade financeira acima da renda necessária para gastos essenciais devam contribuir mais para a manutenção do bem comum do que indivíduos que mal conseguem satisfazer suas urgências básicas. Em geral, o pensamento majoritário converge para a idéia de que todos os indivíduos devem sofrer o ônus da manutenção do Estado e dos bens públicos através de “cotas de sacrifício” iguais para todos, seja através da tributação diferenciada para indivíduos que possuem diferente capacidade financeira de forma

² A revisão de cálculo do PIB realizada pelo IBGE provocou perda de continuidade a partir de 2006 dos dados de arrecadação comparados com o PIB (disponibilizados publicamente pela RFB).

proporcional ao nível de renda (equidade vertical), seja pela imposição de obrigações semelhantes a indivíduos com capacidade econômico-financeira similar (equidade horizontal).

Estes fundamentos dão suporte a políticas que visam diminuir o grau de desigualdade de recursos entre os indivíduos. O imposto de renda pessoa física pode funcionar como meio de redistribuição de recursos, através da tributação da riqueza e posterior transferência do governo por meio de programas sociais visando principalmente os menos favorecidos.

A renda de um determinado indivíduo é percebida pela sociedade como uma medida da capacidade de consumir, e conseqüentemente, uma medida do valor sobre o qual se calculará o ônus que recairá sobre aquele indivíduo. A tributação sobre a renda dos indivíduos é, portanto, intimamente ligada a considerações sobre equidade; é talvez o tributo em que este conceito seja mais valorizado.

O processo de apuração da base de cálculo do imposto consiste em determinar a capacidade contributiva em função da renda auferida, e adequá-la de acordo com as características particulares de cada indivíduo. A distribuição mais equitativa do encargo de pagar o imposto através de uma base de cálculo individualizada e progressiva em relação ao poder aquisitivo é considerada uma grande virtude do imposto de renda, mas também a principal razão da complexidade e do alto custo administrativo desta modalidade de tributo.

As regras gerais que disciplinam a qualificação dos contribuintes, determinação da renda, apuração e pagamento o imposto de renda sobre a pessoa física no Brasil guardam semelhanças com critérios adotados em outros países, como Estados Unidos, por exemplo.

Segundo a definição constante do Regulamento do Imposto de Renda, o contribuinte do imposto de renda pessoa física é o indivíduo domiciliado ou residente no Brasil titular da

disponibilidade de renda e proventos de qualquer natureza³.

A lei pode atribuir à fonte pagadora da renda a obrigação de recolher o imposto (diretamente a partir do contracheque do empregado, por exemplo). A retenção do imposto não confere à fonte pagadora a condição de contribuinte do imposto, mas a torna juridicamente responsável pelo seu recolhimento. O contribuinte real, o indivíduo, tem o dever de recolher o restante do imposto devido, apurado através de ajuste anual por meio de declaração.

Considerando a motivação social que estimula a busca por parte dos governos por um tratamento tributário equitativo para os contribuintes, seria razoável supor que a base tributária integrasse todas as rendas auferidas, independente de sua origem, tal e qual a denominação do tributo (imposto sobre renda e proventos de qualquer natureza). Entretanto, não é o que ocorre. No Brasil, a apuração do total dos rendimentos segue uma lógica similar ao regime de caixa, ao contrário do regime de competência utilizado pelas pessoas jurídicas⁴. Receitas não realizadas não são consideradas “renda” pelo critério adotado no imposto de renda pessoa física.

De acordo com a legislação tributária brasileira, também não são incluídos entre os rendimentos totais (“renda bruta”) benefícios regulamentados por lei e oferecidos pelos empregadores aos empregados, tais como vales transporte, tickets alimentação, planos de saúde e previdência, condicionados a critérios definidos em lei, e benefícios custeados pelo governo a determinadas categorias de servidores públicos, como militares, por exemplo, que podem incluir ajuda de custo, moradia gratuita, etc. Doações recebidas também estão excluídas do conceito de rendimentos brutos

³ Pela lei tributária brasileira, os indivíduos também podem exercer atividades econômicas sob a forma de firmas e empresas individuais que estão excluídas da base de contribuintes do imposto de renda pessoa física, pois são equiparadas a pessoa jurídica.

⁴ Profissionais que exercem atividades sem vínculo empregatício, bem como titulares de cartório e produtores rurais, não precisam observar o regime de competência, e informam seus rendimentos na declaração anual de ajuste em um formulário que representa um “livro caixa” simplificado.

utilizado pela legislação brasileira.

Do ponto de vista econômico, a renda tributável é calculada a partir da renda bruta, através da redução do valor original por meio de deduções, abatimentos e isenções permitidas por lei.

A criação de favorecimentos tributários que reduzem a renda do indivíduo para efeito de apuração do imposto devido pode estar fundamentada em diferentes motivações. Em primeiro lugar, pelo já citado interesse em proporcionar tratamento equitativo entre os contribuintes. Quando o Estado, por exemplo, autoriza o abatimento de despesas médicas, está claramente admitindo que contribuintes com poder aquisitivo idêntico, porém, um deles saudável e o outro padecendo de grave enfermidade, não estão em patamares equivalentes. O raciocínio é idêntico para o caso de deduções por número de dependentes.

Em segundo lugar, tratamentos tributários diferenciados encontram respaldo nas metas sociais perseguidas pelo governo. As deduções por despesas com educação podem ser entendidas como uma forma de subsídio do Estado a um item considerado estratégico ou de interesse social⁵. Em terceiro, algumas deduções podem ser consideradas uma espécie de ajuste contábil, não endossando uma política social específica. É o caso, na legislação brasileira, das despesas do livro caixa, dedutíveis para que a renda do indivíduo que possui atividade profissional autônoma seja tributado pela renda líquida; e dos valores pagos a título de pensão alimentícia, que são tributados em conjunto com os rendimentos declarados pelo detentor da guarda do alimentando.

Um importante item entre as deduções autorizadas pelo governo é o desconto padrão de uma fração de renda, efetivado por meio de declarações de ajuste simplificadas. O desconto padrão é elaborado principalmente para atingir um universo de contribuintes

⁵ O mesmo ocorre com deduções de fomento a iniciativas vinculadas ao Estatuto da Criança e do Adolescente, e as deduções autorizadas para financiamento a atividades culturais e audio visuais.

de menor poder aquisitivo. Justifica-se por questões de economia: administrativa, pela desoneração da máquina pública em fiscalizar deduções itemizadas de contribuintes de menor renda, um processo dispendioso em termos de tempo e recursos; e para o contribuinte, uma vez que representa desoneração de custos de *compliance*, através de um procedimento simplificado.

Além das deduções itemizadas e da dedução simplificada, é autorizado o abatimento de um valor fixo anual por dependente declarado pelo contribuinte.

No imposto de renda pessoa física, o próprio contribuinte declara seus rendimentos, apura o valor a pagar e o recolhe aos cofres públicos, deduzindo a importância já recolhida pela fonte pagadora (este processo de compensação pode gerar crédito a ser restituído ao contribuinte). É impraticável para a administração pública tomar parte ativa neste processo, calculando a quantidade de imposto devida por todos os contribuintes. Sendo assim, o procedimento de identificação da renda tributável e cálculo do imposto devido segue a sistemática do lançamento por homologação. Durante um período de tempo após o ano em que a renda foi auferida (no Brasil, 5 anos), o contribuinte está sujeito a auditorias feitas pelo fisco; expirando o prazo legal, o cálculo do imposto devido é tacitamente aceito pela administração pública. Ao fisco, entre outras atribuições, cabe introduzir um fator de risco representado pela quantidade de auditorias fiscais realizadas e pelas penalidades aplicadas ao contribuinte em situação irregular.

2.3 Alíquotas e Progressividade

Há uma permanente polêmica entre setores da sociedade envolvendo a quantidade ideal de alíquotas de imposto de renda pessoa física, bem como seus valores. Há os que advogam pelo aumento do número de alíquotas, bem como dos valores das alíquotas máximas. Há os que defendem o experimento de uma alíquota única, ou “flat tax”, reduzindo a quantidade de isenções.

A rigor, pode-se ter progressividade no imposto através da imposição uma única

alíquota marginal e um limite de isenção. As alíquotas marginais são dispostas em faixas de renda por uma questão de simplicidade. Seria impraticável estabelecer alíquotas médias através de uma função, mesmo que linear, da renda tributável. A progressividade desejada pode ser obtida pelo estabelecimento de alíquotas marginais divididas por faixas de renda, gradativamente aumentadas da menor para a maior faixa.

Com o passar do tempo, as faixas de renda que delimitam a aplicação das alíquotas marginais sofrem corrosão inflacionária. Se os valores inferiores e superiores que delimitam as faixas de renda relativas a cada alíquota não são corrigidos, em termos reais, em valores compatíveis com a inflação observada, as faixas de renda são alteradas para menos. Esta ocorrência é denominada “draga fiscal” (em inglês, “*bracket creep*”)⁶.

Os contribuintes são obrigados a informar, na declaração de ajuste anual do imposto de renda – pessoa física as rendas provenientes de ganhos de capital, juros, aluguéis, dividendos, ganhos com operações no mercado financeiro. As alíquotas incidentes sobre estas modalidades de renda são em geral diferenciadas, com exceção dos dividendos recebidos, que não sofrem tributação na pessoa física.

Entretanto, convém ressaltar que a tributação sobre o capital é um problema essencialmente dinâmico, envolvendo preços e consumo em um modelo intertemporal. O modelo aqui utilizado para descrever o imposto sobre a renda do trabalho é estático; a tributação sobre o capital foge ao escopo do trabalho⁷.

⁶ No Brasil, entre 1996 e 2001 a tabela de limites de aplicação alíquotas do imposto de renda – pessoa física foi congelada. Foram feitos reajustes esporádicos em 2002 e 2005, e a partir de 2007 até 2010, a tabela foi corrigida com base no programa de metas de inflação. Durante o intervalo de tempo entre 1996 e 2001, portanto, o país assistiu a alteração dos limites de alíquota em termos reais.

⁷ Por outro lado, nem sempre a divisão entre rendimentos do trabalho e do capital é tão clara: tome-se como exemplo os rendimentos recebidos por um médico no exercício de sua profissão. Eventualmente, buscando uma tributação favorável através de planejamento tributário, o médico opta por receber seus rendimentos através de uma pessoa jurídica; mas o pagamento recebido continua a ser rendimento

2.4 Participação Relativa do Imposto de Renda – Pessoa Física no Sistema Tributário Nacional

Alguns fatores aparentemente servem de estímulo contra o aumento da participação relativa do imposto de renda sobre pessoas físicas frente ao total da carga tributária brasileira. As regras constitucionais de repartição de receita, tornaram mais vantajoso política e financeiramente para o poder público federal buscar financiamento através da instituição de novos tributos (particularmente contribuições) do que a partir da reformulação de impostos já existentes.

O nível de renda relativamente baixo, e o alto grau de desigualdade de renda existente na sociedade brasileira contribuem para que a maior parte da população brasileira esteja excluída da base de incidência do imposto de renda – pessoa física. Em 2005, para uma população economicamente ativa de mais de 80 milhões, apenas pouco mais de 7 milhões pagaram imposto de renda. Tal fato concorre para que seja registrado um aumento significativo da base de contribuintes que pagam IRPF sempre que a tabela de faixas de renda por alíquota não é corrigida através dos índices de inflação. Immervoll et al (2010), através de um modelo de microsimulações, mostraram que a tributação sobre a renda das pessoas físicas no Brasil é consideravelmente sensível à chamada “draga fiscal”.

A arrecadação do imposto de renda – pessoa física também sofre com o alto grau de informalidade da economia brasileira. Notoriamente boa parte das receitas advindas da economia informal são evadidas e terminam por não sofrer a incidência da tributação.

decorrente do trabalho. A situação envolve uma mudança de comportamento induzida pela maior tributação sobre a renda da pessoa física em relação à tributação sobre a pessoa jurídica, ou seja, faz parte da distorção econômica provocada pela tributação da renda sobre o trabalho. Porém, se o mesmo médico adquirir um determinado equipamento e utilizá-lo nos diagnósticos, a renda decorrente da utilização do equipamento é economicamente caracterizada não como renda do trabalho, mas como rendimento sobre o capital.

Na faixa de renda mais alta a evasão fiscal dá lugar predominantemente ao planejamento tributário. A título de exemplo, no Brasil a carga tributária sobre empresas que desfrutam de regimes eletivos de tributação como o SIMPLES e o Lucro Presumido acaba sendo mais vantajosa em relação à tributação sobre a renda das pessoas físicas, o que torna compensador para profissionais liberais e autônomos pertencentes a um nível de renda mais alto receber rendimentos através de pessoa jurídica.

Já na fronteira da evasão fiscal, proprietários e sócios de pessoas jurídicas utilizam o expediente de adquirir bens patrimoniais, ativos financeiros e patrocinar gastos e despesas não corporativos com cartões de crédito e viagens através da empresa.

Atualmente os grandes contribuintes do imposto de renda – pessoa física são os assalariados regularmente empregados. Uma evidência é o volume de recursos recolhidos aos cofres públicos através da retenção de imposto de renda direto da fonte pagadora (folha salarial), bem superior ao montante recolhido através do ajuste anual.

3 CUSTOS DE EFICIENCIA E A ELASTICIDADE DA OFERTA DE TRABALHO

3.1 Tributação Sobre a Renda e os Custos de Eficiência

Um sistema econômico é dito eficiente no sentido de Pareto, se não existe possibilidade de melhorar o nível de bem estar de um indivíduo sem que algum outro indivíduo piore em termos de bem estar. O conceito formulado através do primeiro teorema do bem estar garante a eficiência da alocação de recursos na economia em um mercado competitivo, mas não garante a satisfação dos indivíduos com o resultado final. Nos mais diversos grupos sociais, podemos perceber que existe em comum a aversão à desigualdade entre indivíduos. O grau de aversão pode ser grande ou pequeno, dependendo da sociedade. Porém o fato é que em praticamente todos os grupos sociais, é reprovável a idéia de que uma pessoa possa ser dona de todas as riquezas materiais, enquanto outra seja absolutamente carente de bens e serviços, o que dentro do conceito de Pareto-ótimo, seria uma alocação eficiente de recursos.

Entretanto, de acordo com o segundo teorema do bem estar, o governante que atua como planejador de políticas públicas pode redistribuir a dotação inicial entre os indivíduos para alcançar um ponto eficiente no sentido de Pareto que seja socialmente desejável. Para isso bastaria ao governo aplicar tributos e subsídios *lump sum* sobre os indivíduos, e garantir que o mercado competitivo faça o restante do serviço. Se o governo aplica um sistema tributário *lump sum* em um ambiente de competição perfeita, não é possível substituí-lo por outro sistema tributário em que haja melhora do nível de bem estar de um indivíduo sem que outro piore em termos de bem estar.

Na prática, porém, esta solução é impossível. O governo precisaria tributar cada indivíduo em função de suas habilidades naturais, como meio de redistribuição de dotações; o problema é que o formulador da política tributária não detém informações que mostrem todas as características intrínsecas de cada indivíduo, sobretudo as que não possam ser distorcidas ou modificadas, e que seriam necessárias à cobrança eficiente de um imposto. Estas informações são de conhecimento exclusivo do próprio indivíduo, que não tem nenhum incentivo em revelá-las. Em consequência, o governo

somente pode impor tributos sobre características observáveis, como a renda ou o consumo, por exemplo, e que invariavelmente estão sujeitas à mudança de comportamento do indivíduo quando este percebe que de alguma forma pode obter vantagem econômica, maximizando sua utilidade sob a forma de redução do ônus de pagar imposto.

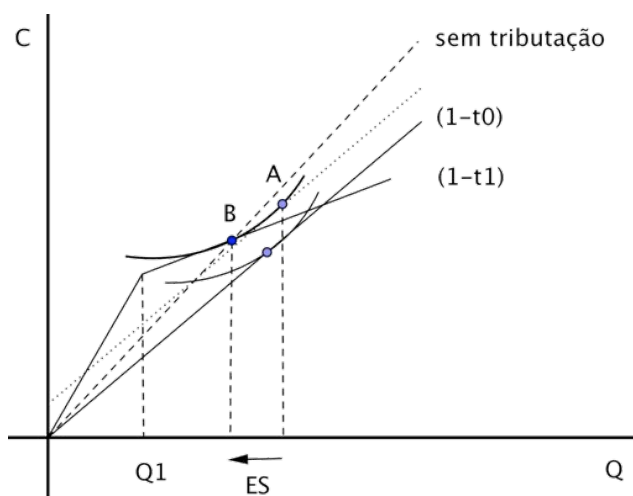
Outra alternativa de efetivação de um sistema de tributação *lump sum* seria aplicar o tributo sobre características completamente alheias a qualquer critério econômico, como a altura das pessoas, por exemplo. Mas tal imposição causaria aversão geral e extrema graças ao sentido de justiça partilhado pela sociedade, e representam um custo político altíssimo aos implementadores do sistema. Sua aplicação é inviável em democracias, e de difícil implantação até mesmo em regimes autocráticos.

É evidente a necessidade do governo em exigir tributos para manter bens públicos, desenvolver programas de interesse social e redistribuir renda. Considerando a inevitabilidade do emprego de tributos que distorcem o comportamento do indivíduo, saber a dimensão do custo arcado pela sociedade pela imposição de um tributo distorcivo é uma informação valiosa para a elaboração, planejamento e execução de políticas públicas. Um método de avaliação e mensuração do custo relacionado à tributação distorciva (o montante que é perdido pela sociedade por causa da transferência de recursos dos indivíduos para o governo) é a utilização do conceito de perda de peso morto (*deadweight loss*).

Conceitualmente, a perda de peso morto é a diferença entre o valor que seria arrecadado caso a tributação fosse *lump sum* e o valor do imposto arrecadado pelo governo. O valor arrecadado por uma tributação *lump sum* é precisamente o quanto o indivíduo pagaria, ou o quanto seria dele retirado para que os preços dos bens fossem conservados em seus patamares anteriores, ou seja, sem tributação, o que coincide com a definição de variação equivalente, uma das medidas de variação do bem estar da economia. No apêndice A um exemplo gráfico ilustra as relações entre os conceitos de variação equivalente, perda de peso morto, efeito renda e efeito substituição.

É a mudança na alíquota marginal que gera o efeito substituição, a perda de peso morto, e o custo de eficiência da economia. A alíquota média pode ser alterada pelo fornecimento, pelo governo, de uma renda não condicionada à oferta de trabalho a todos os indivíduos; neste caso somente observaríamos o efeito renda. Como resultado, podemos ter um sistema tributário onde contribuintes que pagam zero de alíquota média geram custo de eficiência para a economia.

A figura a seguir é um exemplo desta afirmativa. Um modelo de tributação linear é substituído por um sistema tributário em que, primeiramente, os contribuintes usufruem de um programa de subsídios (por exemplo, um imposto de renda negativo), e a partir da renda Q_1 são submetidos à tributação sobre a renda a uma alíquota superior à do modelo linear ($(1 - t_1) < (1 - t_0)$). Embora B esteja na linha de 45 graus que representa tributação zero na média, a mudança marginal causa o efeito substituição (ES) representado pela mudança de A para B, com reflexos em termos de custo de eficiência para a economia.



Para os casos de tributação sobre a renda do trabalho, a medida da variação equivalente possui algumas particularidades. Conforme observado em Creedy (2000), a variação equivalente e variação compensatória, dentro da avaliação de decisões de trabalho e lazer face a uma tributação sobre a renda envolvem também a variação do parâmetro renda (diferentemente do que ocorre com a tributação de commodities).

Considere $C = (\bar{L} - l)w$, onde C é o consumo do indivíduo (tratado como numerário), \bar{L} a quantidade total de horas disponíveis para o indivíduo distribuir entre trabalho e lazer (l), e w a taxa de remuneração por hora de trabalho (valor correspondente à produtividade ou qualificação individual). Podemos especificar o problema da minimização da função gasto do consumidor como

$$\min_{C,l} C + lw$$

s.a

$$U(C,l) = \bar{U}$$

onde o preço do consumo é numerário. As demandas hicksianas (compensadas) definidas pelo problema são respectivamente $C_H(w,U)$ e $l_H(w,U)$.

Consideremos que $m = \bar{L}w$, sendo m a renda total do indivíduo, e \bar{L} a quantidade total de horas disponíveis para o indivíduo distribuir entre trabalho e lazer. Se pudesse trabalhar o tempo todo, o indivíduo ganharia m . Ao optar por uma hora de lazer o indivíduo deixa de ganhar a taxa de remuneração w . Portanto, w é o “preço” de uma hora de lazer. Entretanto, w diminui com a imposição de uma tributação sobre a renda a uma alíquota marginal t , ou seja, $w_1 = w_0(1 - t_1)$. Quando isso acontece, $m_0 \neq m_1$. Partindo das funções gasto $e(w_0, v(w_0, m_0 - EV)) = m_0 - EV$ e $e(w_0, v(w_1, m_1)) = m_0 - EV$, onde $v(w, m)$ é a função utilidade indireta, e EV a variação equivalente, somando e subtraindo m_1 , a variação equivalente passa a ser

$$EV = e(w_1, v(w_1, m_1)) - e(w_0, v(w_1, m_1)) + (m_0 - m_1)$$

em que $(m_0 - m_1)$ representa um ajuste decorrente do aumento do tributo (e consequente redução da renda total após a tributação). Entretanto, Atkinson e Stiglitz (1980) fornecem uma definição alternativa da função gasto que se torna mais adequada ao cálculo da variação equivalente, em que $e(w, U)$ passa a ser relacionada com uma renda lump sum, e não com o gasto total:

$$\min_{C,L} C - Lw$$

s.a

$$U(C,L) = \bar{U}$$

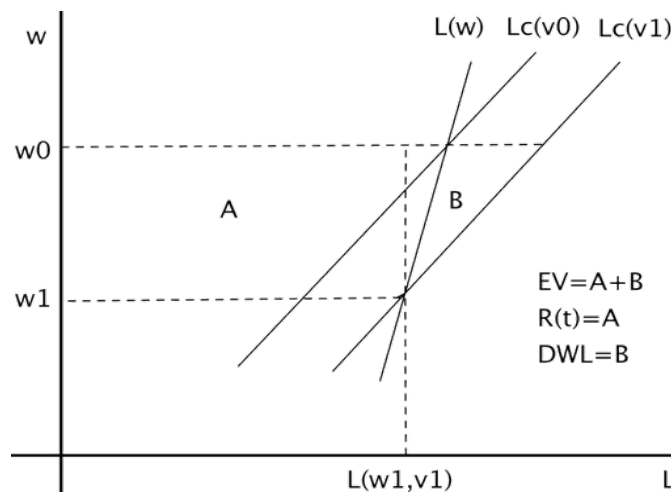
Neste caso, como L equivale a horas de trabalho, e $L = \bar{L} - l$, e $e(w,U) = C_H(w,U) - \bar{L}w + l_H(w,U)$, então a variação equivalente passa a ser

$$EV = e(w_1, v(w_1, m_1)) - e(w_0, v(w_1, m_1))$$

definida em função da oferta compensada (hicksiana) de trabalho. Integrando e aplicando o lema de Shephard, temos:

$$\int_{w_1}^{w_0} L_C(w, v(w, m_1)) dw,$$

que corresponde à área delimitada pelos pontos w_0 , w_1 e pela curva de oferta de trabalho compensada $L_C(v_1)$ conforme ilustrado na figura 5 a seguir:



A figura, bem como a última fórmula, ilustra a medida da variação equivalente, resultante da aplicação de uma alíquota de imposto sobre a taxa de remuneração w . Na mesma figura, $L(w)$ representa a curva marshalliana, a área do retângulo A representa a receita tributária obtida e a área do triângulo B corresponde à perda de peso morto (DWL).

Através da diferença entre a variação equivalente (EV) e a receita tributária ($R(t)$) obtemos a medida da perda de peso morto, ou o valor a mais que poderíamos recolher de tributo além da quantidade efetivamente recolhida, se a tributação fosse *lump sum*⁸. Ou seja,

$$DWL = EV - R(t) = EV - (w_0 - w_1)L(w_1, v(w_1, m_1))$$

A perda de peso morto pode ser mensurada com o uso de elasticidades, a partir da equação de Slutsky. Quanto maior a elasticidade compensada, maior a resposta gerada pelo contribuinte, em reação ao aumento da tributação. O método tradicional empregado para estimar a perda de peso morto causada pela tributação fez uso da elasticidade da oferta de trabalho, que veio a se tornar parâmetro chave para avaliar os efeitos decorrentes de políticas tributárias incidentes sobre a renda a partir do custo de eficiência da tributação. É o que veremos a seguir.

3.2 Tributação Sobre a Renda e Oferta De Trabalho

3.2.1 Modelo Básico

A análise dos efeitos da tributação sobre a renda sobre a oferta de trabalho utiliza o modelo básico da teoria da demanda do consumidor. No modelo, o consumidor possui

⁸ Uma aproximação bastante utilizada para medir a perda de peso morto, seguindo o trabalho de Harberger (1964), é considerar que, de forma simplificada, graficamente a área da perda de peso morto resulta em um triângulo, tal e qual a figura 3. A diferença entre w_0 e $w_1 = w_0(1 - t)$ é $w_0 t$. A imposição do tributo provoca uma variação da oferta do trabalho. A área do triângulo é $1/2(wt)\Delta L$; $\Delta w = wt$, logo a expressão da área do triângulo pode ser rescrita como $1/2(wt)^2 \Delta L / \Delta w$. Considerando que a elasticidade compensada (hicksiana) da oferta de trabalho em relação à taxa de remuneração por hora trabalhada é $\varepsilon^h = \frac{w}{L} \frac{\partial L}{\partial w} \bigg|_U$, então $\partial L \approx (L/w) \varepsilon^h \partial w$, e a perda de peso morto é $DWL \approx \frac{Lw}{2} \varepsilon^h t^2$.

Da equação anterior podemos tirar algumas importantes conclusões. Em primeiro lugar, o resultado mostra a magnitude da evolução do custo de eficiência medido através da perda de peso morto em função da variação do tributo sobre a renda: pela fórmula de Harberger um aumento na alíquota do imposto faz com que o custo de eficiência da tributação tenha um aumento igual ao quadrado do aumento da alíquota. Em segundo lugar, o custo de eficiência é proporcional à elasticidade da oferta de trabalho compensada.

um determinado conjunto de preferências entre consumo e lazer (tomando lazer como o total de horas em que o indivíduo não está trabalhando). O consumidor tem uma quantidade fixa de horas para dividir entre trabalho e lazer; oferta horas de trabalho para auferir renda e consumir. Entretanto, o trabalho é um “mal” e gera desutilidade ($\partial U / \partial L < 0$). O consumo do indivíduo pode ser tratado como um bem composto hicksiano que representa a soma dos n bens consumidos pelo indivíduo, ponderada pelos respectivos preços ($\sum_{i=1}^n p_i q_i$); este consumo é financiado por uma renda total composta por renda advinda do trabalho (Lw), e renda proveniente de outras fontes (I). Considerando t um tributo a uma alíquota fixa, que incide de forma proporcional sobre a renda do trabalho, o problema do consumidor é maximizar a função utilidade $U(C, L)$, sujeito à restrição de consumo $C = Lw(1 - t) + I$. L é ofertado à taxa de remuneração w . Usualmente, em um modelo de equilíbrio estático de oferta do trabalho, a quantidade de horas trabalhadas é utilizada como medida da oferta de trabalho, indicando a quantidade de trabalho ofertada pelos indivíduos.

No problema do consumidor, a inclinação da restrição orçamentária corresponde ao valor da renda obtida do trabalho após a incidência do tributo, $w_0 = w(1 - t_0)$. O intercepto é o ponto I_0 , que equivale a renda do indivíduo que não depende do aumento da quantidade de trabalho oferecida. Entretanto, o problema pode ser apresentado de outra forma. A renda bruta antes do imposto é $Q = Lw$; consequentemente, sendo o total de horas trabalhadas $L = Q/w$, o argumento da função utilidade pode ser modificado, e U pode ser reescrita como $U(C, Q/w)$. Em uma representação gráfica plotando consumo (C) e renda bruta antes do imposto (Q), a inclinação da restrição orçamentária passa a ser simplesmente $(1 - t_0)$, e não mais é afetada por alterações na taxa de remuneração w .

Se a alíquota do tributo é modificada para $t_1 > t_0$ a inclinação da restrição orçamentária diminui, girando em torno do ponto I_0 conforme a figura 6:

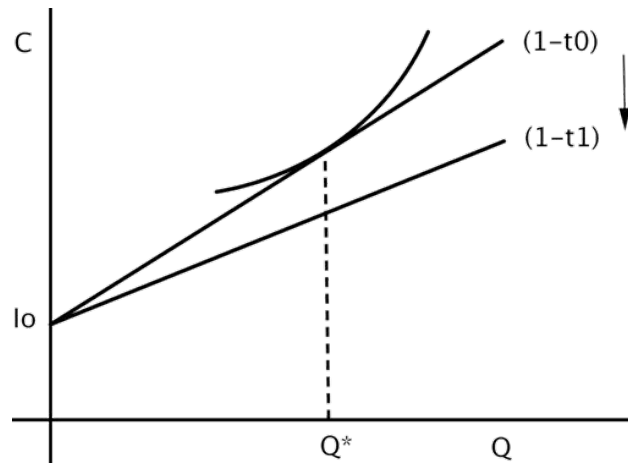


Figura 6

O contribuinte terá uma perda de receita, que refletirá em um nível de utilidade inferior ao que estava anteriormente. Porém os efeitos do aumento do tributo sobre a quantidade de horas trabalhadas não são tão evidentes. Derivando a equação de Slutsky em termos de oferta do trabalho (Apêndice B), os efeitos renda e substituição são dispostos em termos separados. A partir das condições de primeira ordem do problema de maximização da função utilidade, obtemos as funções marshallianas $L = L^*(w, I)$ e $C = C^*(w, I)$. Analogamente, pelo problema dual, minimizando I extraímos as funções hicksianas $L = L^H(w_0, U)$ e $C = C^H(w_0, U)$.

Através da identidade $L^H(w_0, U) = L^*(w_0, I(w_0, U))$, e utilizando o teorema do envelope, a equação de Slutsky escrita em termos de oferta do trabalho é

$$\left. \frac{\partial L}{\partial w_0} \right|_I = \left. \frac{\partial L}{\partial w_0} \right|_U + L \frac{\partial L}{\partial I} \quad (1)$$

O primeiro termo do lado direito da equação é o termo compensado, responsável pelo efeito substituição; o segundo termo é o efeito renda, ponderado pela quantidade de trabalho (em horas) ofertada. Se \bar{L} é o tempo máximo em um dia a ser dividido entre trabalho e lazer (excluindo das 24 horas do dia o tempo que o indivíduo passa dormindo), o tempo gasto com lazer é $l = \bar{L} - L$. Considerando lazer um bem normal, temos que $\partial l / \partial I > 0$, e $\partial L / \partial I < 0$. Já no termo compensado, para o bem demandado

(lazer), $\partial l / \partial w_0|_U < 0$, e $\partial L / \partial w_0|_U > 0$. Equivale dizer que não é possível, com base na equação de Slutsky, determinar o sinal decorrente da soma dos efeitos renda e substituição no modelo de oferta de trabalho proposto; a resposta em termos de variação de horas de trabalho a um aumento de tributação é indefinida, sendo necessária a obtenção de resultados empíricos que esclareçam se a soma dos efeitos renda e substituição representará um aumento ou uma redução na quantidade de horas de trabalho ofertadas.

3.2.2 Modelo Não Linear

O modelo apresentado no item anterior, com restrição orçamentária linear e alíquota do tributo sobre a renda do trabalho constante, é uma simplificação bastante afastada da realidade. Um modelo mais condizente com o mundo real leva em consideração que em geral, tributos incidentes sobre a renda são não lineares, determinando que a restrição orçamentária do modelo seja não linear. Usualmente nos sistemas tradicionais de tributação sobre a renda, alíquotas de imposto são atribuídas a faixas de renda específicas, o que determina que as curvas de restrição orçamentária sejam segmentadas (*piecewise*). O ponto que define a transição entre a alíquota anterior e a seguinte é angulado (“*kink point*”); a função não é bem comportada, e em consequência, não pode ser diferenciada para se obter a dimensão do efeito compensado e efeito renda através da equação de Slutsky.

O imposto de renda sobre as pessoas físicas, via de regra, segue o critério de progressividade de alíquotas; este padrão é em geral definido em função do anseio da sociedade por uma distribuição equitativa da carga tributária.

Considere um sistema de tributação sobre a renda do trabalho, consistindo de duas alíquotas marginais e uma faixa de isenção. O argumento da função utilidade é descrito em termos de renda bruta antes do imposto, ou $U = U(C, Q/w)$. Até Q_1 , a alíquota marginal t_0 é igual a zero; a partir de Q_1 a alíquota é t_1 ; e de Q_2 em diante, alíquota passa a ser t_2 . A tributação é progressiva, ou seja, $t_2 > t_1$, o que faz com que a restrição orçamentária dividida em três segmentos de reta, e tenha formato convexo. O

segmento S_0 tem inclinação de 45 graus ($C = Q$); S_1 e S_2 possuem inclinação de $(1 - t_1)$ e $(1 - t_2)$ respectivamente, conforme mostrado na figura 7 a seguir:

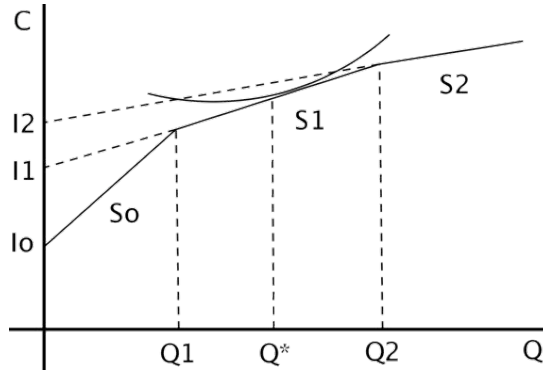


Figura 7

Na figura acima, cada segmento de reta que compõe a restrição orçamentária pode ser estendido até o eixo das ordenadas. O ponto encontrado pode ser interpretado como a renda não proveniente do trabalho “ajustada”, chamado também renda virtual. Hausman (1985), utilizou o conceito para analisar a restrição orçamentária do problema de tributação não linear, e avaliar os efeitos renda e substituição correspondentes. Um indivíduo, cuja quantidade ótima de trabalho ofertada recaia no segmento de reta onde vigora a alíquota marginal t_2 (cujo coeficiente angular é $w(1 - t_2)$), teria uma renda virtual de I_2 . No “kink point” L_2 , os segmentos S_1 e S_2 se encontram, então

$$\begin{aligned} I_1 + (1 - t_1)wL_2 &= I_2 + (1 - t_2)wL_2 \\ I_2 &= I_1 + (t_2 - t_1)wL_2 \end{aligned} \quad (2a)$$

De maneira similar, $I_1 = I_0 + (t_1 - t_0)wL_1$. Desta forma podemos reescrever I_2 como

$$I_2 = I_0 + (t_1 - t_0)wL_1 + (t_2 - t_1)wL_2 \quad (2b)$$

A equação acima mostra que a renda virtual I_2 depende da renda não proveniente do trabalho I_0 ; entretanto, depende também das alíquotas t_0 e t_1 que compõem os coeficientes angulares de S_0 e S_1 , respectivamente. A renda virtual I_2 varia não apenas em função da alíquota t_2 , mas também em função das alíquotas t_0 e t_1 , referentes aos segmentos S_0 e S_1 , que correspondem as faixas de renda inferiores.

Caso, em função das preferências do indivíduo, a quantidade ótima de trabalho ofertado em termos de horas L^* permaneça no mesmo segmento de reta após uma alteração de alíquota tributária neste segmento, ou em faixas de renda inferiores (e mantendo-se a progressividade de alíquotas e a convexidade da restrição orçamentária), o efeito da variação pode ser analisado utilizando a equação de Slutsky. Entretanto, a renda virtual I_i , ao contrário da renda não proveniente do trabalho I_0 , é função da alíquota do tributo. Assim, sendo t a variável exógena do sistema, $w_0 = w_0(t)$ e $I = I(t)$, e tomando a equação de Slutsky em termos de oferta de trabalho, tem-se a seguinte diferenciação, considerando os efeitos indiretos da variável exógena:

$$\frac{dL}{dt} = \frac{\partial L}{\partial w_0} \bigg|_U \frac{dw_0}{dt} + L \frac{\partial L}{\partial I} + \frac{\partial L}{\partial I} \frac{dI}{dt} \quad (3)$$

O efeito total sobre as horas de trabalho em função da variação da renda (por conta de uma variação em t) é indefinido, porém a dimensão do efeito renda é alterada; a renda virtual $I(t)$ faz com que o efeito renda seja maior.

Supondo conhecidas a renda do trabalho antes da tributação w , e as preferências do indivíduo, um sistema tributário como o descrito compreende as alíquotas t_0 , t_1 e t_2 e as quantidades de trabalho em termos de horas ofertadas L_1 e L_2 . A variação simultânea de todos esses parâmetros dentro do modelo pode mostrar indefinições quanto aos efeitos da tributação sobre o trabalho somente passíveis de algum esclarecimento através da análise empírica. No entanto, alguns exemplos simples baseados na variação de um dos parâmetros acima podem facilitar a compreensão desses efeitos. Nos exemplos abaixo considera-se que lazer é um bem normal:

1. Variação da alíquota t do segmento correspondente a faixa de renda acima daquela em que se situa a renda bruta antes do imposto auferida pelo indivíduo Q^* (figura 8). Não há qualquer efeito sobre o indivíduo, e conseqüentemente, não há mudanças em Q^* .

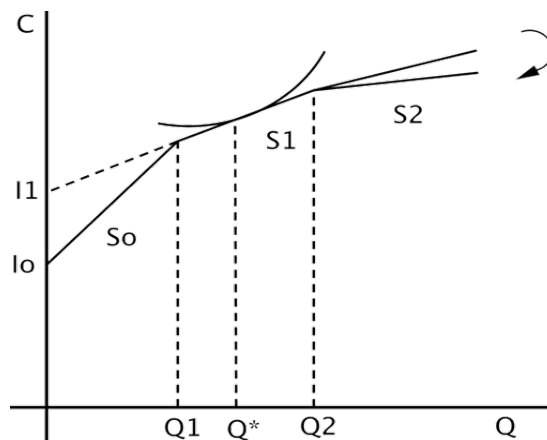


Figura 8

2. Variação de t no segmento que contém Q^* . Com a variação da alíquota marginal do imposto (e consequentemente de $(1 - t)$), o coeficiente angular do segmento S_1 e a renda virtual I_1 variam; há, portanto, a presença dos efeitos substituição e renda. Todavia, tomando como exemplo uma majoração na alíquota marginal t_1 , o somatório dos efeitos não permite concluir o indivíduo ofertará mais ou menos horas de trabalho; consequentemente não se pode afirmar se a variação na renda bruta antes da tributação Q^* será positiva ou negativa.

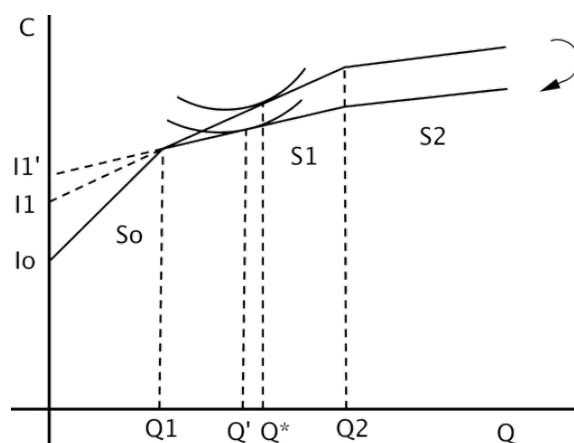


Figura 9

3. Variação de t no segmento abaixo daquele que contém Q^* . A renda virtual varia, porém a inclinação do segmento permanece a mesma. Consequentemente, somente se observa o efeito renda. Se lazer é um bem normal, um aumento em

t no segmento inferior reduz a renda virtual, aumenta a quantidade de horas de trabalho ofertadas pelo indivíduo e sua renda bruta antes do imposto.

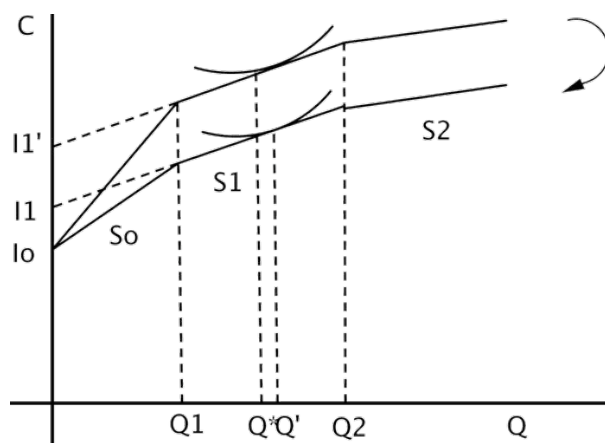


Figura 10

4. Q^* está situado na angulação entre dois segmentos ("kink point") e varia a alíquota do segmento acima ou abaixo de Q^* . A variação de alíquota não provoca qualquer mudança de comportamento do indivíduo em relação a quantidade de horas de trabalho ofertadas.

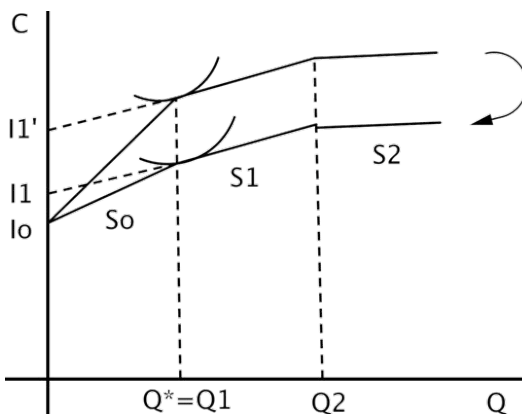


Figura 11

5. Variação nos limites da faixa de incidência referentes às alíquotas marginais t_1 e t_2 . Para os indivíduos acima do limite, a inclinação do segmento permanece a mesma, logo não há efeito substituição. Já a renda virtual é modificada, e há efeito renda. No caso de aumento do limite, o efeito renda é positivo. Em uma

situação em que o valor nominal do limite Q_1 é corroído pela inflação (redução de Q_1 em termos reais), o efeito renda é negativo, e o resultado esperado seria o aumento da quantidade de horas de trabalho ofertadas pelo indivíduo. Já para os indivíduos situados entre o limite antigo Q_1 e o novo limite em termos reais Q_1' , na região em que houve aumento da alíquota marginal do imposto, haveria também efeito substituição, e o sinal resultante da soma dos efeitos seria indeterminado, considerando lazer como um bem comum.

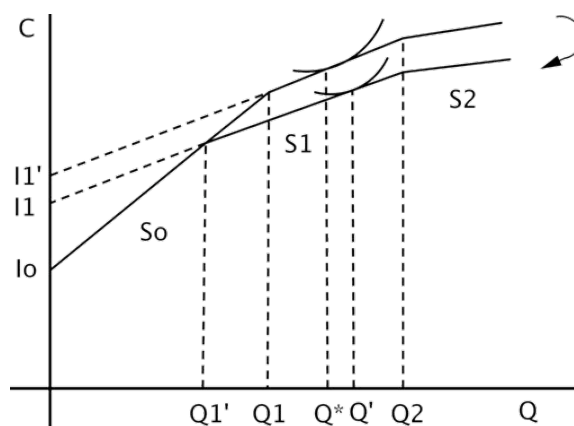


Figura 12

3.2.3 Modelo Não Convexo

Frequentemente, na execução de suas políticas públicas, os governos recorrem a programas de transferência de renda e seguridade social, destinados a classes sociais mais pobres, como ferramentas de distribuição de renda e redução de desigualdades. A saída do indivíduo destes programas, em função de um acréscimo de renda proveniente do trabalho, produz efeito equivalente à tributação sobre a renda. Aplicado ao modelo tradicional de demanda por consumo e lazer, o sistema tributário não linear deixa também de ser convexo. Com a não convexidade, as curvas de indiferença passam a tangenciar em mais de um ponto a restrição orçamentária. Desta forma, passa a existir a possibilidade de mais de uma solução ótima para a quantidade de horas trabalhadas. Além disso, uma pequena variação em algum parâmetro do sistema tributário pode ter como resultado uma grande variação na quantidade de horas de trabalho ofertadas. Nesse caso, a equação de Slutsky não mais pode ser utilizada. Em

consequencia, é difícil obter uma avaliação de efeitos com base exclusivamente no modelo teórico, sem o respaldo de uma análise empírica. A figura 13 exibe um caso em que a não convexidade resulta em mais de uma solução ótima em função da existência de mais de um ponto de tangência entre as preferências do indivíduo e a restrição orçamentária:

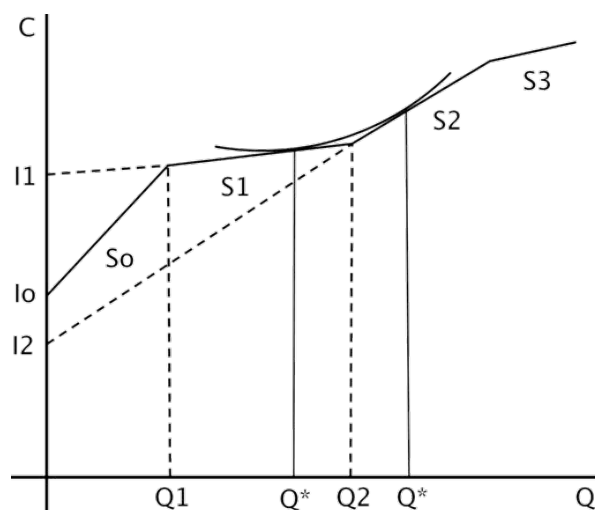


Figura 13

A figura 14 ilustra um modelo com restrição orçamentária não convexa em que, a partir do patamar de renda antes do imposto Q_1 , a saída gradativa do programa de transferência de renda faz com que de fato o indivíduo seja exposto a uma alíquota marginal efetiva mais alta (embora a alíquota definida pela tabela regulamentada pelo sistema tributário seja baixa ou mesmo zero). Ao cessar todos os efeitos do programa, a alíquota passa a ser a regularmente definida pelo sistema tributário. A não convexidade faz com que haja um salto na renda antes do imposto, decorrente da variação de quantidades de horas a serem ofertadas pelo indivíduo:

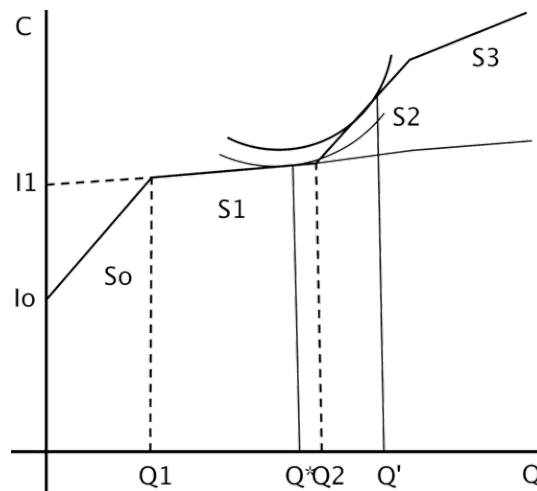


Figura 14

Os exemplos relacionados anteriormente mostram o efeito gerado por uma mudança na alíquota do imposto de renda pessoa física sobre a quantidade de horas de trabalho oferecidas pelo indivíduo. No entanto, na grande maioria dos casos, a autonomia do indivíduo é limitada e ele não pode escolher trabalhar duas horas por dia, por exemplo. A quantidade de horas de trabalho pode ser determinada pela empresa que contrata, por acordos ou pela lei trabalhista. Há custos suportados pelo indivíduo (transporte por exemplo) que restringem as opções a um valor mínimo aceitável de salário e de horas trabalhadas. A decisão que leva uma segunda pessoa da família (a esposa, por exemplo) ao mercado de trabalho pode influenciada por uma mudança de política tributária; outras decisões como o momento da aposentadoria também sofrem influência da tributação. Todas estas situações dizem respeito à decisão de participar ou não do mercado de trabalho, ou seja, a um efeito extensivo na oferta de trabalho provocado pela tributação.

3.2.4 Equação de Slutsky Para a Oferta De Trabalho

O emprego de elasticidades é o formato usual para analisar o efeito causado por uma mudança na tributação na quantidade de horas de trabalho ofertada e a sensibilidade da oferta de trabalho a esta variação. Partindo da equação de Slutsky em (1), obtemos

$$\left. \frac{w_0}{L} \frac{\partial L}{\partial w_0} \right|_I = \left. \frac{w_0}{L} \frac{\partial L}{\partial w_0} \right|_U + w_0 \frac{\partial L}{\partial I} \quad (4)$$

O termo a esquerda da equação é a elasticidade não compensada da oferta de trabalho (marshalliana) ε^m , o primeiro termo a direita da equação é a elasticidade compensada (hicksiana) ε^h , e o segundo termo é um parâmetro η associado ao efeito renda. Segue que a equação de Slutsky em termos de elasticidades é

$$\varepsilon^h = \varepsilon^m - \eta \quad (5)$$

A elasticidade compensada ε^h é derivada do termo da equação de Slutsky responsável pelo efeito substituição, e consiste no parâmetro de maior interesse. Ele mede a variação da oferta de trabalho em resposta a uma alteração da renda do trabalho paga por hora (ou a variação marginal da alíquota do imposto incidente sobre a renda), mantendo a utilidade constante. Como não é diretamente observável, é obtida a partir dos demais parâmetros da equação.

O parâmetro η , relaciona a renda por hora de trabalho w_0 com a variação da renda virtual I . Como a variação de I tem o efeito similar a uma tributação lump sum, e por consequência, não distorciva, a elasticidade compensada da oferta de trabalho ε^h pode ser utilizada como parâmetro indicador da eficiência de uma política tributária sobre a oferta de trabalho. Quanto mais sensível o indivíduo ao aumento ou diminuição de imposto, maior a elasticidade compensada ε^h .

O parâmetro ε^h é utilizado na avaliação do custo de eficiência do imposto de renda sobre a economia causado por uma mudança tributária (*deadweight loss*, ou perda de peso morto).

No processo de arrecadação de tributos, o governo arrecada dos indivíduos os tributos, que em seguida, são restituídos à sociedade. Com a tributação, o indivíduo deixa de ganhar, pois abre mão de horas de trabalho por conta de uma remuneração por hora (w) inferior; o governo, por sua vez, deixa de arrecadar sobre este salário. A consequência é a perda de eficiência que penaliza a economia como um todo.

3.2.5 Estimativas de Elasticidade da Oferta de Trabalho

Uma extensa literatura aborda a estimativa dos parâmetros de elasticidade não compensada da oferta de trabalho e da elasticidade renda, a partir dos quais se obtém a elasticidade compensada da oferta de trabalho⁹.

Embora diferentes métodos de estimativa tenham sido empregados, focalizando grupos diversos (trabalhadores homens, trabalhadoras mulheres, mães solteiras, trabalhadores por faixa de renda, etc), com resultados por vezes divergentes, é possível generalizar algumas conclusões com base nos resultados obtidos. Os resultados mais relevantes foram compilados por Meghir e Phillips (2010)¹⁰. Os resultados sugerem que a força de trabalho masculina apresenta uma elasticidade compensada bastante baixa, muito próxima de zero. Mulheres, particularmente mães solteiras e integrantes de famílias que já possuem fonte de renda, apresentam elasticidade um pouco maior. Entretanto, de forma geral, os resultados sinalizam que a mudança de alíquota do imposto sobre a

⁹ Uma representação econométrica de um modelo de tributação linear, em que o indivíduo é remunerado por hora de trabalho a taxa de $w_0 = w(1 - t_0)$ e a função marshalliana da oferta de trabalho é $L = L^*(w, I)$, pode ser descrita como

$$L_i = \beta_0 + \beta_1 w_{0i} + \beta_2 I + \beta_3 X_i + \varepsilon_i$$

Onde, para o indivíduo i , w_{0i} é a remuneração por hora de trabalho livre do imposto t_0 , I é a renda não laboral, X_i é um vetor que reúne características específicas do indivíduo, β_1 é a elasticidade não compensada da oferta de trabalho e β_2 é o parâmetro referente ao efeito renda. A estimativa pode ser feita utilizando o método de mínimos quadrados ordinários. Porém, modelos de tributação não linear requerem o uso de metodologia específica, como a desenvolvida por Hausman (1985).

¹⁰ Em Meghir e Phillips (2010) há uma extensa série de tabelas, que listam uma grande variedade de resultados de elasticidades compensadas da oferta de trabalho, obtidos a partir de estudos empíricos. A título de ilustração, para indivíduos do sexo masculino, que compõem a maior parte da oferta de mão de obra, MaCurdy, Green e Paarsch (Estados Unidos, 1990), encontraram um valor de -0,01, Blomquist e Newey (Suécia, 2002) estimaram em -0,02 e Bourguignon e Magnac (França, 1990) em -0,07.

renda não altera de forma sensível a quantidade de horas trabalhadas pelos indivíduos, o que sugeriria também que a tributação sobre a renda não seria muito distorciva. Em termos de bem estar, a tributação geraria uma perda de peso morto pequena.

Vários aspectos dificultam a obtenção de estimativas de elasticidades da oferta de trabalho com base em mudanças de políticas tributárias, e consequentemente, a avaliação destas políticas considerando bem estar, variação de receitas obtidas pelo governo, e sensibilidade relativa dos diversos grupos de indivíduos face uma alteração da carga tributária.

O aperfeiçoamento dos métodos econométricos visando a estimativas a partir de modelos não lineares tem sido a alternativa para contornar dificuldades como a não linearidade em função de alíquotas de imposto progressivas ou de programas sociais e de distribuição de renda, que provocam a endogeneidade das variáveis w (renda recebida pelo indivíduo por hora de trabalho) e I (renda não laboral).

Já a utilização de modelos de escolha discreta, obtendo elasticidades próprias tem sido a forma utilizada para lidar com a decisão de participação do mercado de trabalho encarada pelo indivíduo, motivada por custos inerentes a própria atividade, como transporte, alimentação, gastos com “*baby sitters*” no caso de mães solteiras, etc.

As firmas usualmente oferecem postos de trabalho exigindo uma quantidade fixa de horas, não permitindo, ao menos no curto prazo, que o indivíduo tenha margem para optar por uma determinada quantidade de horas entre um leque contínuo de opções.

A maior aptidão para o trabalho de determinados indivíduos que por isso possuem maior qualificação técnica e maiores salários produz um viés, pois esta variável, não observada, integra o termo erro, e é correlacionada com w . O problema é parcialmente corrigido através da utilização de dummies com características do indivíduo, e variáveis instrumentais.

3.2.6 A Elasticidade da Oferta de Trabalho - Limitações Ao Estimar Custos de Eficiência

Modelos de tributação ótima e estimativas de perda de peso morto que traduzem os custos de eficiência provocados por mudanças em políticas fiscais utilizaram por muito tempo a elasticidade da oferta de trabalho como parâmetro que refletia o comportamento dos indivíduos diante das alterações de alíquotas marginais de imposto. Um aspecto que assume grande importância diz respeito ao uso das horas de trabalho do indivíduo como medida de quantidade de trabalho empregada pelos modelos de oferta de trabalho. O esforço de trabalho não consiste unicamente na disposição de trabalhar mais ou menos horas por dia, por semana ou por mês. A intensidade com que o indivíduo realiza o trabalho, a energia, a criatividade a dedicação ou a concentração empregadas, não podem ser medidas em horas. Tampouco é trivial medir as horas de trabalho de um indivíduo que possui um trabalho autônomo, sem horários fixos ou dias de folga definidos. Um indivíduo com a motivação adequada pode trabalhar mais e com mais qualidade, em um tempo menor que o previsto. É evidente que, considerando o esforço ou intensidade com que o trabalho é realizado, não podemos mais falar de uma renda bruta do trabalho por hora (w) fixada exogenamente, pois o comportamento do indivíduo interfere no valor de w .

Estas considerações sugerem que os valores encontrados para a elasticidade da oferta de trabalho, em termo de variação de horas trabalhadas, subestimam a reação do indivíduo à política tributária, bem como o efeito substituição e os custos de eficiência em termos de bem estar decorrentes. Sugerem também a busca de um parâmetro mais abrangente, capaz de incorporar tais reações.

Além disso, existência de uma base tributária complexa, com isenções e deduções, faz com que o indivíduo, ao procurar consumir bens favorecidos pela regra tributária, cause distorções e gere ineficiências na economia.

Abordada em dois papers seminais de Feldstein (1995, 1999), a utilização da elasticidade da renda tributável (ou ETI, de *elasticity of taxable income*) como parâmetro abrangente, capaz de incorporar o efeito da tributação sobre o esforço

despendido no trabalho (e não horas) e o efeito substituição decorrente da alteração de preços relativos entre formas de consumo com tratamento tributário mais ou menos favorecido, abriu uma ampla fronteira de pesquisa ainda em desenvolvimento.

4 A ELASTICIDADE DA RENDA TRIBUTAVEL (ETI)

4.1 Modelo Básico

Desde a análise de Harberger (1964) sobre os reflexos da imposição de tributos, em termos de eficiência, sobre a economia, a elasticidade sobre a oferta de trabalho passou a ser utilizada como parâmetro chave para avaliar as distorções causadas pela carga tributária suportada pelos indivíduos. Entretanto, vimos que as horas trabalhadas são uma *proxy* inexata para a quantidade de trabalho realizada, pois esta envolve outras dimensões como, por exemplo, a intensidade de esforço realizado ou a complexidade do trabalho que o indivíduo aceita executar.

Feldstein (1995, 1999) enfatiza que este é apenas um impacto a levar em consideração no caso de uma mudança na política fiscal que altere as alíquotas marginais dos tributos incidentes sobre a renda dos indivíduos. Estes podem maximizar sua renda de diversas outras formas; contribuintes podem modificar a forma como consomem, como são remunerados, como declaram seus rendimentos, como poupam ou investem, optando por alternativas menos onerosas para si em termos tributários.

Essa vasta gama de reações era relevada pelos economistas até então, principalmente porque incluir cada reação observada em um único modelo gera um problema de alta complexidade e difícil construção. Feldstein (1995, 1999) propõe uma alternativa de grande simplicidade e elegância, ao sugerir a utilização de um parâmetro único incorporando todas estas reações: a elasticidade da renda tributável.

A elasticidade da oferta de trabalho mede a variação da quantidade de trabalho oferecida pelo indivíduo, medida em horas, em relação a variação de renda auferida por hora (w_n), composta por uma taxa fixa w multiplicado por pelo fator de renda líquida após o imposto ($1 - t_n$). Já a ETI mede a variação na renda tributável Q , em relação a uma variação no fator de renda líquida após o imposto ($1 - t$).

O clássico problema do consumidor para a oferta de trabalho é o ponto de partida para o modelo de renda tributável de Feldstein (1999). Diante um conjunto de preferências

entre consumo e lazer, o indivíduo maximiza sua utilidade, sujeito a uma restrição orçamentária ao consumo:

$$\begin{aligned} & \max_{C,L} U(C,L) \\ & \text{s.a} \\ & C = Lw(1-t) + I \end{aligned}$$

Tomando o total \bar{L} de horas disponíveis a serem gastas entre trabalho e lazer como numerário, o total L de horas trabalhadas é igual a $(1-l)$, onde l é lazer. A renda bruta é $Q = Lw$. Feldstein argumenta que, além de lazer, o indivíduo consome benefícios que são oferecidos como compensações, e que não compõem a renda tributável. Além disso, outras formas de consumo são isentas, e por isso, são deduzidas da base cálculo do imposto. Q passa a ser a renda tributável e é definida como

$$Q = w - (wl + E + B) \quad (6)$$

onde B é o total de benefícios e compensações oferecidos ao indivíduo, e que não estão incluídos na renda do trabalho oferecida a tributação; e E refere-se aos itens de consumo isentos, e deduzidos da base de cálculo do imposto sobre a renda. A restrição orçamentária passa a ser

$$\begin{aligned} & \max_{C,Q} U(C,Q) \\ & \text{s.a} \\ & C = (1-t)Q + I \end{aligned} \quad (7)$$

As condições de primeira ordem $U_C(C,Q) = \lambda$ e $U_Q(C,Q) = (1-t)\lambda$ levam à função marshalliana $Q = Q^*(1-t, I)$. A minimização de I , sujeito a $U(C,Q) \geq \bar{U}$ resulta em $Q = Q^H(1-t, U)$ e na equação de Slutsky em termos de renda tributável

$$\left. \frac{\partial Q}{\partial (1-t)} \right|_I = \left. \frac{\partial Q}{\partial (1-t)} \right|_U + Q \frac{\partial Q}{\partial I} \quad (8)$$

uma pequena manipulação algébrica conduz à equação de Slutsky em elasticidades

$$\xi^h = \xi^m - \eta \quad (9)$$

onde $\xi^m = \frac{(1-t)}{Q} \frac{\partial Q}{\partial(1-t)} \Big|_I$ é a elasticidade não compensada da renda tributável, e

$\xi^h = \frac{(1-t)}{Q} \frac{\partial Q}{\partial(1-t)} \Big|_U$ é a elasticidade compensada da renda tributável, o parâmetro de

maior interesse. O parâmetro $\eta = (1-t) \frac{\partial Q}{\partial I}$, é referente ao efeito renda, e pode ser

transformado para $\eta = \kappa(Q(1-t)/I)$, onde $\kappa = \frac{I}{Q} \frac{\partial Q}{\partial I}$, é a elasticidade de Q em relação a

renda virtual. Se lazer, benefícios e isenções forem bens normais, Um aumento da renda virtual gera um aumento no consumo destes bens, e κ terá sinal negativo.

Através da elasticidade compensada da renda tributável ξ^h , tem-se um parâmetro de medida mais abrangente do efeito substituição causado pela variação do imposto. Isto porque a elasticidade da renda tributável capta variações em E e B, tais como benefícios concedidos ao indivíduo, ou isenções usufruídas, mesmo que L permaneça constante. Além disso, a “renda tributável” agrega mais informações, e por isso aparenta ser uma *proxy* mais razoável do esforço de trabalho do que “horas trabalhadas”. Em consequência, a elasticidade da renda tributável pode medir também as diversas dimensões da atividade laboral, tais como intensidade, criatividade, complexidade ou o trabalho executado em *home office*.

Um aumento de imposto (t), faz com que $(1-t)$ diminua. Havendo efeito substituição, um aumento em t e a correspondente diminuição de $(1-t)$ induz o indivíduo a procurar pelo consumo de bens e atividades menos sujeitas a tributação, o que também reduz a renda tributável Q (ou seja, $\partial Q / \partial(1-t) > 0$ e ξ^h tem sinal positivo). Neste caso um aumento do tributo provoca o aumento da perda de peso morto e um aumento na arrecadação em uma proporção menor que o acréscimo de imposto. Caso a ETI seja muito alta, pode ocorrer até mesmo a diminuição na arrecadação do imposto.

O que resulta desta diferença conceitual entre a elasticidade da oferta de trabalho e a elasticidade da renda tributável?

Embora haja números discordantes, a maior parte das pesquisas efetivadas para medir a elasticidade da oferta do trabalho chegou a resultados muito pequenos, próximos ou iguais a zero para homens, e um pouco maiores para mulheres, o que sugeriria que a economia perderia pouco ou nada em termos de eficiência pelo aumento da carga tributária. Neste caso, os cofres públicos recolheriam integralmente o valor correspondente ao aumento, e que o indivíduo nada faria além de trabalhar como sempre e pagar na totalidade o aumento de imposto. Todavia a percepção geral diante dos resultados das reformas tributárias empreendidas por diversos países sinaliza que não é isso que ocorre. Outras mudanças comportamentais contribuem juntamente com a variação na quantidade de trabalho para gerar uma perda de peso morto, e consequentemente, maior ineficiência na economia. Possivelmente, estes outros fatores têm um peso muito maior que a variação em horas da quantidade de trabalho na geração de toda esta distorção. Em um modelo empírico de avaliação de efeitos da tributação sobre a oferta do trabalho, estes fatores acabam integrando o termo erro.

Se incorporamos individualmente cada um destes efeitos em um modelo, através de parâmetros estruturais específicos, temos que dotar este modelo de tal estrutura, que o trabalho empírico torna-se extremamente difícil. Além disto, o mais relevante é poder avaliar os efeitos comportamentais de uma mudança na política tributária, obtendo indicadores de eficiência e arrecadação que avaliem a reforma como um todo. Assim, torna-se conveniente o uso de um parâmetro com a capacidade de integrar todos estes efeitos.

A elasticidade da renda tributável, por sua simplicidade e a possibilidade de agregar variados efeitos, não importando quais sejam as causas, é atualmente vista como parâmetro chave na avaliação de políticas tributárias. Sob determinadas condições que serão vistas mais adiante, pode vir a ser uma estatística suficiente para obtenção de estimativas confiáveis de custos de eficiência da economia, variações de receitas governamentais, bem como para a composição de modelos de tributação ótima.

4.2 Estudos Empíricos - Resultados Encontrados Na Literatura

A reforma tributária americana de 1986 que reduziu as alíquotas marginais do imposto de renda pago pelos contribuintes americanos, em alguns casos de 50% para menos de 30%, foi analisada por Feldstein (1995,1999), a partir de um painel de dados a contendo uma amostra estratificada de contribuintes cedida pelo Internal Revenue Service (IRS). Os dados empíricos foram tratados pelo método de diferenças em diferenças, comparando dados antes e depois da reforma. A estimativa de Feldstein para a elasticidade da renda tributável mostrou valores superiores a 1, alcançando valores acima de 2 para contribuintes de alto poder aquisitivo, sugerindo grande sensibilidade deste parâmetro à mudança da alíquota de imposto. Quanto maior a elasticidade, maior o custo marginal de eficiência por real arrecadado como consequência do aumento de imposto. Em consequência, a estimativa levou a expressivas importâncias referentes a perda de peso morto e a variação na receita de arrecadação do imposto.

Em estudos posteriores (Auten e Carroll, 1999; Gruber e Saez, 2002; Saez, 2003; Kopczuk, 2005 entre outros) as estimativas foram refeitas, e o resultados obtidos para a elasticidade da renda tributável foram mais modestos, mas ainda assim expressivos. Estes resultados se devem à utilização de tratamentos econométricos visando reduzir os problemas encontrados na obtenção de estimativas. Para os Estados Unidos, a maior parte das estimativas oscila entre 0,3 e 0,5. Embora em diversos países da OCDE existam estudos visando obter estimativas da ETI, em países em desenvolvimento essas estimativas ainda são raras.

Utilizando como base o painel da NBER contendo dados do imposto de renda para contribuintes americanos ao longo dos anos 80, Gruber e Saez (2002) realizaram a estimativa da ETI mais consistente até o momento. A metodologia empregada neste estudo visou reduzir efeitos prejudiciais, que afetam a obtenção de estimativas mais acuradas, caso da reversão à média (causada pelo fato de que alguns indivíduos podem não auferir renda alta de forma constante, e outros apresentam um rápido crescimento de renda em um curto prazo) e de tendências decorrentes do ambiente macroeconômico, como concentração de renda, ou progresso tecnológico afetando um

grupo de indivíduos ao longo do tempo. Em todos estes casos, as mudanças de renda não guardam relação com qualquer mudança de política tributária, e produzem uma correlação espúria, que em geral fazem com que os resultados sejam superestimados.

Gruber e Saez (2002) utilizaram variáveis instrumentais por conta da endogenia causada pela progressividade do tributo, eliminaram contribuintes com renda anual menor que \$ 10.000 para controle da reversão à média, e empregaram dados relativos a um período mais extenso de tempo (1979 a 1990), o que ajuda a controlar efeitos intertemporais, e ao mesmo tempo proporcionar uma avaliação dos efeitos de mais de uma alteração na política tributária americana. Além disso, incluíram um parâmetro de efeito renda, ausente em trabalhos anteriores, e estimaram não só a elasticidade da renda tributável mas também a elasticidade da renda bruta, que possui uma definição mais abrangente e mais próxima da renda total do indivíduo¹¹.

Uma variável dependente representativa do efeito renda foi introduzida no modelo, e os valores pequenos ou estatisticamente insignificantes encontrados motivaram Gruber e Saez (2002) a assumir que a elasticidade compensada e a não compensada são idênticas. A hipótese do efeito renda não significativa foi levantada em *papers* anteriores, e foi mantida na maioria dos *papers* subsequentes.

¹¹ Gruber e Saez (2002) utilizaram um modelo econométrico com a seguinte especificação:

$$\log\left(\frac{Q_{t+i}}{Q_t}\right) = \beta_0 + \beta_1 \log\left(\frac{1-t_{t+i}}{1-t_t}\right) + \beta_2 ano + \beta_3 estcivil + \sum_{i=1}^j \beta_{4i} spline_i + \varepsilon$$

A variável dependente $\log(1-t_{t+i}/1-t_t)$ original é substituída por uma variável instrumental para evitar problemas de endogeneidade (um choque positivo em ε pode alterar $\log(1-t_{t+i}/1-t_t)$, visto que t é progressivo e torna-se maior com o aumento da renda); uma função spline é utilizada para controle de tendências sobre a renda não causadas por tributos; e *estcivil* e *ano* são dummies de ano base e estado civil. O parâmetro β_1 representa a elasticidade da renda tributável.

Gruber e Saez (2002) encontraram um valor médio para a elasticidade compensada da renda tributável em torno de 0,4. O valor da elasticidade da renda bruta ficou em torno de 0,12, o que corrobora a importância das deduções, isenções, benefícios e demais expedientes utilizados para reduzir a base de cálculo do tributo em reação ao valor total da ETI.

As estimativas feitas por Gruber e Saez (2002) evidenciam que a elasticidade compensada da renda tributável varia em função de níveis de renda, e convergem em apontar maiores elasticidades para indivíduos de alta renda. Este resultado é intuitivo, pois indivíduos de grande poder aquisitivo normalmente lançam mão dos mais variados expedientes para evitar a tributação, enquanto contribuintes mais pobres em geral não têm a mesma facilidade. Quanto mais alta a renda, maior o acesso a métodos sofisticados de evitar a tributação.

4.3 Restrições e Limitações

Saez, Slemrod e Giertz (2009) relacionam as mais importantes limitações e restrições que devem ser consideradas na utilização da ETI em análises de bem estar. A maioria diz respeito a externalidades.

Muitas vezes, uma reforma tributária provoca concorrência entre bases tributárias do mesmo sistema. Um aumento do imposto de renda sobre pessoas físicas, por exemplo, pode gerar uma migração de contribuintes para a base de tributação do imposto de renda sobre pessoas jurídicas. Boa parte dos contribuintes, particularmente aqueles de maior poder aquisitivo, tem meios e recursos para exercer suas atividades profissionais sendo tributados sob nova base, com tributação mais favorável. As diferenças apuradas em estimativas decorrentes da mudança de bases não se devem a perdas, mas mera transferência entre bases tributárias. O efeito resultante é o de superestimar as estimativas da elasticidade compensada da renda tributável.

Uma reforma que aumente a carga tributária pode alterar decisões de consumo ao longo do tempo. Em nosso caso, a perspectiva de aumento de imposto pode provocar

decisões que antecipem o recebimento de rendas, visando aproveitar as condições tributárias atuais, e evitando as condições futuras. Goolsbee (2000), analisou o aumento do imposto de renda americano em 1993, antecipado pelos contribuintes, e mostrou que a elasticidade da renda tributável no curto prazo era bem maior que no longo prazo para indivíduos da alta renda, devido principalmente ao exercício de opções de posse destes indivíduos. No caso oposto, uma perspectiva de redução do imposto pode fazer com que os indivíduos posterguem o recebimento de suas rendas. As consequências de uma variação na carga tributária ao longo do tempo atingem decisões de poupança, investimentos e planejamento da carreira profissional. Boa parte destas variações pode ser avaliada como efeito substituição intertemporal, entre curto e longo prazo. Um repentino aumento ou redução da renda tributável, atribuído a efeitos intertemporais deve ser levado em consideração nas estimativas de receitas e eficiência econômica resultantes do uso da elasticidade da renda tributável.

Outra importante constatação acerca da elasticidade compensada da renda tributável, observada por Slemrod e Kopczuk (2002) e Kopczuk (2005) é a de que o parâmetro em si está sujeito a influência da orientação política tributária imprimida pelos governantes. Esta influência pode ser determinada por uma mudança na base tributária, aumentando ou reduzindo a quantidade de itens dedutíveis, ou mudando critérios de auditoria ou penalidades. Desta forma, ao invés de avaliar as alterações comportamentais do indivíduo, a ETI passa a refletir as mudanças de política do governo.

4.4 Elisão e Evasão Fiscal

A elisão e evasão fiscal estão entre as externalidades que afetam o bem estar, mas devem ser analisadas com cuidado no que diz respeito à utilização da ETI como parâmetro de medida do custo de eficiência da tributação. Evasão fiscal é a atuação ilegal do indivíduo visando omitir do fisco parte dos rendimentos recebidos. A elisão é a utilização de brechas e lacunas na legislação para diminuir os rendimentos sujeitos à tributação.

O imposto de renda sobre as pessoas físicas, com sua diversidade de alíquotas, deduções e exclusões estimula a criação dos mais diversos meios de contornar ou burlar a exigência legal do tributo; em geral, quanto mais complexo o sistema tributário, mais sujeito está às práticas de elisão e evasão.

A prática de elisão e evasão produz reflexos na forma de perda de bem estar para a economia. O risco assumido pelo indivíduo ao optar pela prática da sonegação reverte em perda de eficiência econômica. Custos adicionais são suportados pelo contribuinte em ferramentas de elisão fiscal, na forma de contratação de escritórios de advogados na montagem de estruturas de planejamento tributário, advogados, despesas bancárias, abertura de pessoas jurídicas, e outros artifícios custosos que não integrariam as preferências de consumo do indivíduo, se não houvesse a tributação. O mesmo pode ser dito para a utilização de testas de ferro (“laranjas”), montagem de caixa dois e outros expedientes ilegais próprios da evasão fiscal. Desta forma, reflexos decorrentes da evasão e elisão integram o efeito substituição, e constituem uma parte do valor observado da elasticidade compensada da renda tributável.

Na formulação original de Feldstein, a elasticidade da renda tributável agrega diversos comportamentos distintos, decorrentes do conjunto de variáveis que reduzem a renda do indivíduo até o patamar de renda tributável sobre a qual incidirá a alíquota do imposto. Sob uma ótica do bem estar, e considerando por hipótese o aumento da carga tributária, é indiferente se o fator que provocou a redução da renda tributável foi decorrente de uma alteração na oferta de trabalho ou de sonegação.

Entretanto, deve-se atentar para o fato que multas e penalidades pecuniárias cobradas pela administração pública, sob uma perspectiva de bem estar, não são vistas como perdas relacionadas com a evasão e elisão; de fato são transferências de recursos, e se não levados em consideração podem interferir nas estimativas da ETI. Parte destas perdas não representa efetivamente custo social, e sim um custo privado do agente que deu causa; as perdas deste agente são transferidas aos cofres da administração pública sob a forma de multas aplicadas em razão de procedimentos de fiscalização e não representam perda de bem estar para a sociedade. Chetty (2009) derivou um modelo que relaxa hipóteses implícitas no modelo de Feldstein (1999), avaliando custos

de transferência para indivíduos neutros ao risco; o resultado tem implicações importantes quanto à permanência da ETI como estatística suficiente para estimativas de bem estar. Na próxima seção, abordaremos este modelo com mais detalhe.

4.5 Modelo de Chetty – Generalização do Modelo Básico

Como visto, sob determinadas restrições, a elasticidade da renda tributável é uma estatística suficiente para avaliação da eficiência e dos custos de uma reforma tributária. Todavia, não eliminou completamente a necessidade de se considerar mecanismos e o comportamento do indivíduo que geram fatores cuja dimensão não está presente na ETI, sendo necessário alguma estrutura e ajustes no modelo. É o caso da influência política no parâmetro, e de externalidades, como a migração do contribuinte entre bases tributárias, dos efeitos intertemporais, e da evasão fiscal.

Analisando o modelo de Feldstein (1999) frente a algumas destas externalidades, Chetty (2009) argumenta que o modelo parte de um pressuposto implícito; o de que o custo marginal de evitar a tributação deve ser igual à alíquota do tributo em condições ótimas.

Para simplificar a notação, vamos fazer uso de L , trabalho, ao invés de l , o lazer; e tomaremos

Seja $D = E + I$, o total das exclusões, compensações e benefícios utilizados pelo indivíduo para evitar a tributação.¹² Presume-se que sem qualquer custo para evitar a tributação, o indivíduo teria incentivo a informar $D = wL$, o que reduziria a zero a renda tributável Q , e consequentemente, a receita do governo. Sendo assim, há um custo que é função de D . Assim reescrevemos o problema de maximização da utilidade como

¹² Partimos do problema de maximização da função utilidade $U(C, Q)$, sujeita à restrição $C = (1 - t)Q + I$, onde $Q = w - (wl + E + B)$.

$$\begin{aligned}
& \max_{L,D} U(C,L) \\
& s.a \\
& C = (1-t)(wL - D) + I - f(D) + D
\end{aligned} \tag{10}$$

Feldstein define C como o consumo do indivíduo que não é favorecido através de benefícios, exclusões, deduções ou regras tributárias. Acrescentamos D em (10), que passa a registrar o consumo total do indivíduo.

$$\frac{\partial U}{\partial D} = 0 \text{ é condição de primeira ordem.}$$

$$\left. \frac{\partial U}{\partial D} \right|_L = \frac{\partial U}{\partial C} \frac{\partial C}{\partial D} = \frac{\partial U}{\partial C} (t - f'(D)) = 0$$

Se $\partial U / \partial C \neq 0$, então $(t - f'(D)) = 0$, e

$$t = f'(D) \tag{11}$$

Chetty (2009) mostra que, na presença de externalidades, $t = f'(D)$ não mais se sustenta. Neste caso, a elasticidade da renda tributável não pode mais ser vista como estatística suficiente para avaliar o custo de eficiência (perda de peso morto) resultante de uma mudança na tributação sobre a renda.

Tomemos o caso em que os custos são em verdade transferências de recursos para outros agentes econômicos. O caso clássico é o da evasão fiscal, em que o contribuinte deixa de declarar ao fisco parte de sua renda, sob o risco de devolver o valor acrescido de multa (transferência de recursos do indivíduo ao governo).

Seja $z(D,t)$ o custo total associado à evasão fiscal, função da receita omitida D e da alíquota marginal do tributo t . O indivíduo enfrenta a probabilidade $p(D)$ de ser pego, e além de pagar o tributo sobre D pode ser multado em $M(D,t)$. Então $z(D,t) = p(D)(tD + M(D,t))$ e o indivíduo maximiza utilidade da seguinte forma:

$$\begin{aligned}
& \max_{L,D} U(C,L) \\
& s.a \\
& C = (1-t)(wL - D) + I - f(D) + D - z(D,t)
\end{aligned} \tag{12}$$

Da condição de primeira ordem $\partial U / \partial D = 0$, temos que

$$\left. \frac{\partial U}{\partial D} \right|_L = \frac{\partial U}{\partial C} \frac{\partial C}{\partial D} = \frac{\partial U}{\partial C} (t - f'(D) - z'(D)) = 0$$

o que leva a

$$t = f'(D) + z'(D) \quad (13)$$

A equação (13) sintetiza que na presença de externalidades como a transferência de recursos decorrente da evasão fiscal, no ótimo, o custo marginal representado pela tributação de 1 R\$ (um real) não mais é igual ao custo de oportunidade de reduzir a renda tributária; passa a ser equivalente a um somatório do custo de oportunidade e do custo de transferência.

4.6 Avaliação De Bem Estar Utilizando o Modelo De Chetty

Passamos a analisar as implicações do modelo de Chetty (2009) sobre as avaliações dos custos de eficiência em termos de bem estar. Com este propósito, define-se o bem estar social $W(t)$ como a soma das utilidades dos indivíduos e da receita tributária auferida pelo governo. Esta por sua vez agrega a arrecadação do imposto ($tQ = t(wL - D)$), e $z(D, t)$, transferido do indivíduo para o governo na forma de multas. A função utilidade é $U(C, L, D) = C - d(L)$ em que $d(L)$ é a desutilidade do trabalho e $C = (1 - t)(wL - D) + I - f(D) + D - z(D, t)$, conforme definido em (12).

Considerando que o indivíduo maximiza $U(C, L, D)$, e que $C(t)$, $L(t)$ e $D(t)$, sendo t exógeno, a função utilidade indireta é $V(C^*(t), L^*(t), D^*(t), t) = V(t)$. Pelo teorema do envelope, os efeitos indiretos sobre a utilidade podem ser relevados, contando somente o efeito direto do parâmetro exógeno. Ou seja,

$$\frac{dV(t)}{dt} = -(wL - D) - \frac{\partial z}{\partial t}$$

Definindo A como o valor arrecadado pelo governo com impostos e multas, a arrecadação marginal dA/dt é

$$\frac{dA}{dt} = (wL - D) + t \left(\frac{d(wL)}{dt} - \frac{dD}{dt} \right) + \frac{\partial z}{\partial D} \frac{dD}{dt} + \frac{\partial z}{\partial t}$$

A variação marginal da função bem estar W devido à mudança da carga tributária é

$$\frac{dW}{dt} = \frac{dV(t)}{dt} + \frac{dA}{dt}$$

que se torna

$$\frac{dW}{dt} = t \frac{d(wL)}{dt} + \frac{dD}{dt} \left(\frac{\partial z}{\partial D} - t \right)$$

sabemos da condição de primeira ordem (13) que $t = f'(D) + z'(D)$, o nos leva a

$$\frac{dW}{dt} = t \frac{d(wL)}{dt} - \frac{dD}{dt} (f'(D))$$

segue uma breve manipulação algébrica

$$\frac{dW}{dt} = t \left[\frac{d(wL)}{dt} - \frac{dD}{dt} \left(\frac{f'(D)}{t} \right) \right] + \frac{f'(D)}{t} \frac{d(wL)}{dt} - \frac{f'(D)}{t} \frac{d(wL)}{dt}$$

chegamos a

$$\frac{dW}{dt} = t \left[\left(1 - \frac{f'(D)}{t} \right) \frac{d(wL)}{dt} + \left(\frac{f'(D)}{t} \right) \frac{d(wL - D)}{dt} \right] \quad (14)$$

O resultado obtido em (14) é importante. Ele mostra que, quando há um aumento na carga tributária sobre a renda, havendo custos transferidos entre agentes da economia, e custos de oportunidade pelo uso de estratégias visando evitar o pagamento do tributo, a perda de eficiência e bem estar na economia é resultante de uma média ponderada entre a variação na renda tributável ($wL - D$) e a variação na renda bruta (wL). O fator de ponderação $f'(D)/t$ representa a razão entre o custo de oportunidade e o total de custos relacionados à redução do valor tributável ($t = f'(D) + z'(D)$).

O resultado demonstra que, na presença de custos de transferência, a elasticidade da renda tributável não é mais uma estatística suficiente para estimativas de perda de bem estar decorrentes de uma alteração do tributo sobre a renda. Além da renda tributável, a renda bruta e o parâmetro $f'(D)/t$, ligado ao custo de redução do tributo, são necessários ao conhecimento do custo de eficiência da economia. É importante notar que, na ausência de custos de transferência, $f'(D)=1$, e o custo de eficiência total é determinado integralmente pela renda tributável. Se há somente custos de transferência, $f'(D)=0$, e o custo de eficiência passa a ser totalmente determinado pela renda bruta. Podemos expressar a equação em forma de elasticidades:

$$\frac{1}{wL} \frac{dW}{dt} = \frac{t}{dt} \frac{d(wL)}{wL} - \frac{1}{wL} \left(\frac{f'(D)}{t} \right) \frac{d(D)}{dt} \frac{D}{D}$$

$$\frac{dW}{dt} = wL \left[\xi - \left(\frac{f'(D)}{t} \right) \frac{D}{wL} \delta \right] \quad (15)$$

em que ξ é a elasticidade da renda bruta em relação à alíquota de imposto; δ é a elasticidade da redução do valor tributável em relação ao imposto exigido; e o já conhecido parâmetro $\left(\frac{f'(D)}{t} \right)$ é ponderado pelo fator que mede a relação entre a renda que escapou da tributação e a renda bruta do indivíduo.

Gorodnichenko, Martinez-Vazquez e Sabirianova (2009) utilizaram (15) para criar um modelo empírico para avaliar os efeitos da reforma tributária implantada na Rússia em 2001, em que houve uma redução dramática da estrutura progressiva de tributação para um sistema tributário com uma alíquota. Baseados em dados de consumo das famílias, e na renda declarada, eles construíram estimativas para ξ , δ e $f'(D)/t$ e concluíram que a grande variação na arrecadação decorrente da reforma é em sua maior parte explicada pela redução da evasão fiscal, bem mais representativa que os ganhos de produtividade da economia.

5 ESPECIFICAÇÃO DE UM MODELO ILUSTRATIVO DA TRIBUTAÇÃO SOBRE A RENDA NO BRASIL CONSIDERANDO O COMPORTAMENTO DOS INDIVÍDUOS

5.1 Derivação do modelo básico

A utilização da elasticidade da renda tributável para avaliação de políticas tributárias ganhou bastante atenção entre pesquisadores de finanças públicas nos últimos anos. Por sua simplicidade, e por abranger diversas formas de resposta à taxação em um único parâmetro, a elasticidade da renda tributável ganhou crescente interesse como parâmetro de análise do sistema tributário, avaliação de custos envolvidos na arrecadação de receita tributária, e no desenho de um sistema tributário ótimo. No entanto, a ETI não dispensa cuidados necessários ao seu uso.

Conforme constataram Slemrod e Kopczuk (2002), a elasticidade da renda tributável em si não é um parâmetro estrutural, pois não é invariante relativamente à execução de políticas públicas. O agente político pode modificá-la, alterando a base tributária através de mudanças da legislação, ou aumentando o esforço de fiscalização, por exemplo. A elasticidade da renda tributável é variável em horizontes temporais de curto e longo prazo; e não há um valor único de ETI, pois contribuintes de diferentes faixas de renda possuem diferentes facilidades de acesso a recursos, visando escapar da tributação, e diferentes percepções de risco em relação à evasão fiscal.

Portanto, embora a princípio possa ser tomada por uma forma reduzida de avaliação do conjunto de reflexos de uma mudança tributária, a utilização da elasticidade não prescinde da construção de um modelo com alguma estrutura. Temos como exemplo a generalização do problema de Feldstein (1999) realizada por Chetty (2009), que resultou na necessidade de estimar dois outros parâmetros, para casos em que haja presença de evasão fiscal. Chetty (2009b) alerta que para a construção de modelos que gerem estatísticas suficientes utilizáveis em análise de bem estar as premissas do modelo devem ser avaliadas. Chetty (2009), ao avaliar a hipótese de que $t = f'(D)$, constatou que a fórmula de Feldstein (1999) superestima os custos de eficiência da tributação, e que seu relaxamento obriga a estimação de novos parâmetros estruturais.

Exemplificando os conceitos abordados ao longo do texto, vamos definir uma forma funcional para a preferência dos indivíduos e um modelo de distribuição das taxas de remuneração que caracterizam as habilidades ("*skills*") dos indivíduos no Brasil. Através de uma simulação simples com dados agregados oriundos da Secretaria da Receita Federal do Brasil relacionando rendimentos tributáveis e arrecadação do imposto de renda de pessoas físicas (IRPF) por faixa de renda, pretendemos mostrar de que forma podemos obter valores representativos das elasticidades da renda tributável, compensada e não compensada, do efeito renda, calcular uma hipotética alíquota marginal única sobre os contribuintes e efetuar uma estimativa sobre a receita do tributo e a perda de peso morto que seriam observadas caso hipoteticamente a alíquota marginal do imposto tivesse sido majorada.

Um dos recursos mais utilizados nas estimativas de ETI é o emprego de uma função quase linear com elasticidade constante. Grande parte da literatura que trata de elasticidade da renda tributável admite por premissa que o efeito renda é negligenciável, o que facilita bastante a especificação funcional de um modelo de maximização de utilidades e a estimativa dos parâmetros de interesse. Entretanto, a hipótese do efeito renda zero, não obstante alguns resultados como o de Gruber e Saez (em que o parâmetro obtido não foi significativo), é uma simplificação que tem a vantagem de contornar as não linearidades no modelo, e que torna a elasticidade observável (não compensada) igual à não observável (compensada, que é o parâmetro de interesse para análises de bem estar)¹³. Em nosso caso, vamos adotar um modelo que permita a ocorrência do efeito renda, apresentando a não linearidade característica do imposto de renda pessoa física e que admita variações progressivas do parâmetro de elasticidade da renda tributável em função das mudanças de habilidades individuais (w), e consequentemente das faixas de renda (o que estaria de acordo com a percepção geral de que indivíduos com maior qualificação e poder aquisitivo têm maior

¹³ Em um modelo quase linear, sem efeito renda, do tipo $U(C, Q) = C - \frac{w}{1 + 1/\zeta} \left(\frac{Q}{w} \right)^{1 + 1/\zeta}$ a solução decorrente das condições de primeira ordem é $Q = Q_0(1 - t)^\zeta$, em que $w = Q_0$ quando não há tributação, e ζ é a elasticidade da renda tributável. A quase linearidade implica que a ETI é constante.

capacidade de fugir de um aumento do imposto). O ponto de partida será a maximização de uma função utilidade separável entre o consumo C e a renda Q (obtida a partir do esforço L empregado a uma taxa de produtividade w).

$$\max U(C, Q) = \frac{C^{1-\rho}}{1-\rho} - \frac{1}{\gamma} \left(\frac{Q}{w} \right)^\gamma, \quad (16)$$

onde ρ é um parâmetro de aversão relativa ao risco, e γ é um parâmetro de desutilidade do trabalho, utilizado na determinação da elasticidade intertemporal de substituição da oferta de trabalho (elasticidade de Frisch)¹⁴. O problema é sujeito à restrição $C = (1-t)Q + I$, onde I é a renda virtual; a condição de primeira ordem pode ser escrita como

$$[(1-t)Q + I]^{-\rho}(1-t) - \frac{Q^{\gamma-1}}{w^\gamma} = 0, \text{ ou}$$

$$-\rho \ln[(1-t)Q + I] + \ln(1-t) - (\gamma-1) \ln Q + \gamma \ln w = 0.$$

A renda tributável Q , e a alíquota marginal de imposto t são dados disponíveis e a renda virtual I é calculada a partir de (4), bastando conhecer elementos básicos da legislação tributária. Consequentemente, w pode ser obtido a partir da condição de primeira ordem:

$$w = \left[\frac{Q^{(\gamma-1)}((1-t)Q + I)^\rho}{(1-t)} \right]^{1/\gamma}$$

A partir da maximização de (16), e das condições de primeira ordem, com o emprego de alguma álgebra (apresentada no Apêndice C), obtêm-se fórmulas específicas para a elasticidade da renda tributável, elasticidade-renda em relação a w , e o parâmetro η do efeito renda, utilizado na equação de Slutsky (5).

¹⁴ Considerando a função utilidade $U(C, L) = C^{1-\rho}/(1-\rho) - L^\gamma/\gamma$, a elasticidade intertemporal de Frisch é $\left. \frac{\partial L/L}{\partial w/w} \right|_\lambda = 1/(\gamma-1)$.

- Elasticidade da renda em relação a w :

$$\frac{dQ}{dw} \frac{w}{Q} = \frac{\gamma}{(\hat{\rho} + \gamma - 1)} \quad (17),$$

$$\text{onde } \hat{\rho} = \frac{\rho(1-t)Q}{Q(1-t) + I}$$

- Efeito renda

$$\eta = \frac{dQ}{dI} (1-t) = -\frac{\hat{\rho}}{(\hat{\rho} + \gamma - 1)} \quad (18)$$

- Elasticidade da renda tributável:

$$\xi = \frac{dQ}{d(1-t)} \frac{(1-t)}{Q} = \frac{\rho Q(1-t) - Q(1-t) + I}{\rho(1-t)Q + (\gamma - 1)(Q(1-t) + I)} \quad (19)$$

A elasticidade compensada da renda tributável pode, conseqüentemente, ser encontrada através da equação de Slutsky:

$$\left. \frac{dQ}{d(1-t)} \frac{(1-t)}{Q} \right|_U = \frac{dQ}{d(1-t)} \frac{(1-t)}{Q} - \frac{dQ}{dI} (1-t) =$$

$$\xi^h = \xi^m - \eta = \frac{\rho Q(1-t) - Q(1-t) + I}{\rho(1-t)Q + (\gamma - 1)(Q(1-t) + I)} + \frac{\hat{\rho}}{(\hat{\rho} + \gamma - 1)} \quad (20)$$

Considerando as fórmulas aqui desenvolvidas, particularmente (20), que resulta na elasticidade compensada da renda tributável, percebe-se que neste modelo a ETI não é constante. Como é função de $(1-t)$ e de ρ , o governo pode alterá-la, através da adoção de políticas públicas como a mudança da alíquota marginal do imposto, ou do aumento do esforço de fiscalização, o que é captado pelo modelo adotado. A constatação reforça a idéia de que, ao relaxarmos as hipóteses que simplificam o modelo e introduzem limitações, fica evidente que a elasticidade da renda tributável não é um parâmetro estrutural, pois não é invariante à execução de políticas governamentais.

5.2 Estimativa de uma alíquota marginal única para o IRPF

O governo busca maximizar o bem estar da população, entretanto os gastos públicos estão sujeitos a restrição orçamentária – a obtenção de receita suficiente (R) para fazer frente aos dispêndios:

$$R = t \int Lw f(w) dw = t \int Q_0(w) f(w) dw,$$

sendo t a alíquota marginal em um regime de tributação linear sobre a renda. $Q_0(w)$ é a renda antes da tributação. Por simplicidade vamos relevar quaisquer correlações ou influências existentes entre as várias modalidades de tributação empregadas no Brasil, e considerar apenas a necessidade do governo tributar a renda dos indivíduos. Na ausência de tributos, $Q = Q_0$. Em Q_0 , a renda virtual $I = 0$, então

$$-\rho \ln[Q_0] - (\gamma - 1) \ln Q_0 + \gamma \ln w = 0$$

$$(\rho + \gamma - 1) \ln Q_0 = \gamma \ln w$$

$$Q_0 = w^{\gamma/(\rho + \gamma - 1)} \quad (21)$$

A função $f(w)$ é a densidade de distribuição da produtividade do indivíduo, decorrente de sua habilidade ou qualificação. Suponhamos que a necessidade de receita por parte do governo é suprida através da arrecadação total T , referente à tributação sobre a renda dos indivíduos, a ser obtida pela imposição de uma alíquota marginal de imposto sobre a renda original dos indivíduos antes da tributação. Então

$$T = t \int Q_0(w) f(w) dw \quad (22)$$

Conhecendo a arrecadação anual de imposto de renda pessoa física e a renda tributável Q , através da qual se calcula w , e por extensão a renda antes da tributação $Q_0(w)$, podemos ter uma idéia da alíquota de imposto t , se o país utilizasse um modelo de tributação de uma única alíquota.

Para modelar a distribuição de w , foi utilizada a Distribuição de Pareto, usualmente utilizada para descrever a distribuição de renda ou riqueza de indivíduos, e utilizada por

Saez (2004), por sua adequação a faixas mais elevadas de renda.

A distribuição de Pareto acumulada é

$$F(w) = 1 - \left(\frac{k}{w} \right)^\theta$$

e a função densidade de distribuição correspondente é

$$f(w) = \frac{\theta k^\theta}{w^{\theta+1}}$$

onde k é o parâmetro de localização que consiste no menor valor assumido por w . Optamos por utilizar como k o valor da produtividade w (calculado a partir da renda tributável) correspondente ao limite inferior da faixa de menor alíquota da tabela de imposto de renda. É necessário que o parâmetro de forma θ seja conhecido através de estimativa.

Se $k \leq w \leq \infty$, e $\theta > \hat{\gamma}$, $\gamma/(\rho + \gamma - 1) = \hat{\gamma}$

$$\int_k^\infty w^{\gamma/(\rho+\gamma-1)} \frac{\theta k^\theta}{w^{\theta+1}} dw = \frac{\theta(\rho + \gamma - 1)k^{\left[\frac{\gamma}{(\rho+\gamma-1)}\right]}}{[\theta(\rho + \gamma - 1) - \gamma]} = E[Q_0(w)] \quad (23)$$

No Apêndice D, apresentamos de forma mais detalhada o cálculo do valor esperado de $Q_0(w)$.

Para estimar a distribuição de Pareto, foi considerado que

$$P(x > w) = \frac{N}{N^*} = \left(\frac{k}{w} \right)^\theta$$

$$N = \left(\frac{k}{w} \right)^\theta N^*$$

e

$$\log N = \alpha - \theta \log w + \varepsilon. \quad (24)$$

A estimativa de θ pode ser obtida diretamente através de Mínimos Quadrados Ordinários (MQO).

5.3 Efeitos sobre a arrecadação e perda de peso morto

Seguindo nosso modelo ilustrativo, faremos uso das expressões para elasticidades e para o efeito renda obtidas na subseção 5.1 a partir da maximização da função de utilidade (16), e da distribuição de w modelada como uma distribuição de Pareto (subseção 5.2), para derivar um modelo em que seja possível obter resultados que caracterizem os efeitos das respostas comportamentais dos indivíduos em termos de bem estar.

A evidência empírica indica que as reações dos contribuintes do imposto são mais intensas em faixas de renda mais altas. Seguindo o desenvolvimento de Saez (2001), vamos supor que o governo providencie um aumento na maior alíquota marginal da tabela do imposto de renda sobre a pessoa física equivalente a dt . Os dados serão os mesmos utilizados anteriormente, constantes do ano-calendário de 1998.

Na ausência de qualquer reação por parte dos indivíduos, a arrecadação esperada seria puramente um “efeito mecânico” sobre a base que sofre o aumento da alíquota marginal do imposto:

$$dM = (\bar{Q} - Q^*) dt$$

em que Q^* é o limite inferior a partir do qual o governo altera a alíquota do imposto, e \bar{Q} é um valor médio representativo da renda tributável dos contribuintes acima do limite Q^* . Um indivíduo altera sua renda em dQ , em resposta a um aumento da alíquota do imposto em sua faixa de renda. Considerando que $Q = Q(1 - t, I)$, a mudança dQ na renda tributável de um indivíduo é representada por

$$dQ = -\frac{\partial Q}{\partial(1-t)} dt + \frac{\partial Q}{\partial I} dI$$

O aumento dI na renda virtual é igual ao aumento na alíquota dt aplicado sobre o limite inferior da faixa Q^* . Define-se a elasticidade não compensada da renda tributável como

$$\gamma^m = \frac{(1-t)}{Q} \frac{\partial Q}{\partial(1-t)} \Big|_I, \text{ e o parâmetro do efeito renda é } \eta = (1-t) \frac{\partial Q}{\partial I}, \text{ conforme apresentado}$$

anteriormente. Segue que

$$dQ = -\frac{\partial Q}{\partial(1-t)} \frac{Q(1-t)}{Q(1-t)} dt + \frac{\partial Q}{\partial I} \frac{(1-t)}{(1-t)} Q^* dt =$$

$$dQ = -(\gamma^m Q - \eta Q^*) \frac{dt}{(1-t)} \quad (25)$$

Para obter o valor da arrecadação adicional conseguida com o aumento do imposto deve-se consolidar o “efeito mecânico” com a variação na receita ocasionada pela reação do contribuinte dQt , ou seja:

$$dR = dM \left[1 - \left(\frac{\gamma^m Q - \eta Q^*}{Q - Q^*} \right) \left(\frac{t}{(1-t)} \right) \right]$$

Pela equação de Slutsky, a elasticidade compensada da renda tributável $\gamma^h = \gamma^m - \eta$.

Substituindo na equação,

$$dR = dM \left[1 - \left(\frac{\gamma^h Q}{Q - Q^*} + \frac{\eta(Q - Q^*)}{Q - Q^*} \right) \left(\frac{t}{(1-t)} \right) \right] =$$

$$dR = dM \left[1 - \left(\frac{\gamma^h Q}{Q - Q^*} + \eta \right) \left(\frac{t}{(1-t)} \right) \right] \quad (26)$$

Se na faixa de renda analisada Q segue a distribuição de Pareto, então

$$Q = \frac{\int_{Q^*}^{\infty} w^{\gamma/\rho + \gamma - 1} f(w) dw}{\int_{Q^*}^{\infty} f(w) dw}$$

fazendo $\gamma/(\rho + \gamma - 1) = \hat{\gamma}$, e $f(w) = \frac{\theta K^\theta}{w^{\theta+1}}$,

$$Q = \frac{\theta K^\theta w^{\hat{\gamma}-\theta} / (\hat{\gamma} - \theta) \Big|_{Q^*}^{\infty}}{\theta K^\theta / \theta w^\theta \Big|_{Q^*}^{\infty}} = \frac{\theta Q^{*\hat{\gamma}}}{(\theta - \hat{\gamma})}$$

Na formula de dR ,

$$dR = dM \left[1 - \left(\frac{\theta}{\theta + K^{1-\gamma}(\hat{\gamma} - \theta)} \gamma^h + \eta \right) \left(\frac{t}{(1-t)} \right) \right] \quad (27)$$

A expressão acima representa a variação total na receita tributária obtida pelo governo com a mudança da alíquota marginal na faixa de alíquota superior. Dentro desta expressão, o último termo equivale ao custo de eficiência marginal decorrente da mudança de alíquota. É a fração perdida em recursos para que possa ser coletado dR em imposto sobre a renda. A variação na perda de peso morto decorrente da reação comportamental dos indivíduos é o custo de eficiência marginal vezes o “efeito mecânico” dM :

$$dDWL = dM \left[\left(\frac{\theta}{\theta + K^{1-\gamma}(\hat{\gamma} - \theta)} \gamma^h + \eta \right) \left(\frac{t}{(1-t)} \right) \right], \quad (28)$$

enquanto $-dDWL/dM$ é a fração de receita perdida decorrente de reações comportamentais do contribuinte do imposto ($dDWL < 0$).

Constata-se que o governo, que pretendia arrecadar dM somente consegue arrecadar $dR = dM + dDWL$. É óbvio que a sociedade paga R\$ 1,00 para arrecadar $(dM + dDWL)/dM$. Ou, expressando de uma outra forma

$$dM/(dM + dDWL) = dM/dR = 1 - dDWL/dR$$

Esta expressão define quantos reais em perda de utilidade para os indivíduos são necessários para aumentar a arrecadação do imposto em 1 real. Este conceito é conhecido por custo marginal de eficiência dos fundos públicos (MECF). O último termo à direita da equação ($-dDWL/dR$) representa o custo de eficiência submetido à

sociedade por cada real a mais arrecadado, em decorrência do aumento de imposto.

5.4 Resultados

Os cálculos desta seção foram feitos com a extração de dados do Texto para Discussão da Receita Federal intitulado “Análise Econômica da DIRPF 1999”, disponível no *site* da Receita Federal do Brasil. São dados referentes ao imposto de renda pessoa física (IRPF) do exercício de 1999 (ano calendário 1998).

Na declaração de ajuste do imposto de renda – pessoa física (DIRPF) exercício de 1999, vigorava a tabela de faixas de renda anual com as seguintes alíquotas marginais: até R\$ 10.800,00, zero; de R\$ 10.800,00 até R\$ 21.600,00, 15%; acima de R\$ 21.600,00, 27,5%.

Naturalmente, o método aqui empregado padece de restrições, algumas já abordadas neste trabalho; principalmente, decorrentes do uso de dados agregados publicados pela Receita Federal, extratificados em onze faixas de renda, correspondentes a um só período temporal (exercício de 1999).

No estudo da Receita Federal os dados disponibilizados utilizam a mesma definição de rendimento tributável existente nas declarações de ajuste do imposto de renda (DIRPF). Ocorre que esta definição exclui as deduções simplificadas e itemizadas, que somente são descontadas dos rendimentos quando é apurada a base de cálculo do tributo. Trata-se de uma classificação característica do Direito Tributário, e diz respeito à incidência ou não do fato gerador do tributo sobre a renda dos indivíduos, antes das deduções regulamentares¹⁵.

Para o nosso propósito, é interessante que estas deduções sejam consideradas. A definição de renda tributável aqui utilizada exclui tais deduções. Aliás, a opção por

¹⁵ O Regulamento do Imposto de Renda define “rendimentos tributáveis” como “...os rendimentos provenientes do trabalho assalariado, as remunerações por trabalho prestado no exercício de empregos, cargos e funções, e quaisquer proventos ou vantagens percebidos...”

deduzir mais despesas da renda potencialmente sujeita ao pagamento do imposto é uma das reações mais evidentes à imposição de uma alíquota maior de imposto, e foi expressamente considerada no estudo de Feldstein (1999).

No estudo da Receita Federal são informados também os valores agregados por faixa de renda das deduções permitidas pela legislação. Estes dados foram utilizados para construir um valor percentual de deduções por faixa de renda (Tabela 1). A renda tributável utilizada para estimativas foi obtida pela subtração deste percentual do valor identificado como “rendimento tributável” na DIRPF (e que corresponde à definição legal, mas não à econômica do termo “rendimento tributável”).

A renda virtual foi calculada tendo em vista a presença do efeito renda no modelo, utilizando a formulação deduzida em (2a) e (2b):

- Faixa 1 (até R\$ 10.800,00) – zero
- Faixa 2 (até R\$ 21.600,00)

$$I_1 + Q(1 - t_1) = Q$$

$$I_1 = Qt$$

- Faixa 3 (acima de R\$ 21.600)

$$I_1 + Q(1 - t_1) = I_2 + Q(1 - t_2)$$

$$I_2 = I_1 + Q(t_2 - t_1)$$

onde I_i é a renda virtual correspondente a cada faixa e t_i a alíquota marginal da mesma faixa.

Os valores dos parâmetros ρ e γ poderiam ser obtidos através da calibração do modelo a partir de dados da economia do Brasil em 1998. Entretanto nosso propósito é bem mais simples. Vamos nos espelhar em valores razoáveis levando em consideração a literatura existente. Chetty (2006) relaciona em tabela estimativas de ρ em estudos sobre oferta de trabalho, e constata que em raros trabalhos as estimativas ultrapassam 1.25; estimativas de γ vão desde valores inferiores em torno de 0.5 para estimativas feitas através do modelo clássico de oferta do trabalho a valores próximos de 2 em

estudos macroeconômicos. Nos cálculos aqui realizados, o valor utilizado para o parâmetro de aversão relativa ao risco ρ foi de 1.1. O parâmetro de desutilidade do trabalho é vinculado à elasticidade intertemporal do trabalho (elasticidade de Frisch). Os valores estimados da elasticidade de Frisch encontrados na literatura variam bastante em função do perfil do contribuinte e método de estimação. Meghir e Phillips (2010) relacionam vários trabalhos onde encontram-se valores em uma gama que vai de 0.1 a 3.8. Utilizaremos o valor encontrado em Kimball e Shapiro (2008) que avalia a dimensão dos efeitos renda e substituição na oferta de trabalho americana, e corresponde à unidade. Com esse valor, o parâmetro de desutilidade do trabalho γ é igual a 2.

Com estes valores, a elasticidade não compensada da renda tributável e a elasticidade compensada da renda tributável oscilaram em torno de 0,098 e 0,594, respectivamente, valores em patamar próximo ao encontrado na literatura, e um pouco acima das estimativas do estudo de Gruber e Saez.

Para estimar o parâmetro θ , foi utilizada a quantidade de contribuintes pertencentes a cada faixa de renda para construir dados de frequência acumulada (N), conforme informações retiradas do histograma publicado no texto da Receita Federal “Análise Econômica da DIRPF 1999”.

Na tabela 1, Q é o limite inferior de cada faixa de rendimento tributável; F é a quantidade de contribuintes declarantes do imposto em cada faixa. N é o quantitativo de contribuintes que auferiram rendimentos acima de Q ; w foi calculado a partir da renda tributável Q :

Tabela1

<i>Faixa de renda (mil R\$)</i>	<i>Deduções por faixa</i>	<i>Q (limite inferior de faixa)</i>	<i>W (limite inferior de faixa)</i>	<i>F</i>	<i>N</i>
0→12	21.73%	0,00	0,00	4.777.068	11.056.038
12→18	25.70%	9.392,27	16.293,35	2.652.812	6.278.970
18→24	32.19%	13.373,16	22.953,07	1.138.450	3.626.158
24→30	31.08%	16.273,65	30.874,55	691.647	2.487.708
30→36	30.15%	20.676,77	38.438,00	444.685	1.796.061
36→48	28.12%	25.146,97	46.165,86	529.960	1.351.376
48→60	24.53%	34.501,04	62.487,80	294.578	821.416
60→72	22.39%	45.283,53	81.523,36	174.606	526.838
72→96	20.09%	55.879,65	100.423,49	170.496	352.232
96→120	17.76%	76.712,74	138.037,14	78.517	181.736
Acima de 120	14.19%	98.682,26	178.226,46	103.219	103.219

Para elaborar a estimativa, foram consideradas as faixas de renda acima de R\$ 12.000,00. Nem todos os contribuintes com renda abaixo deste valor estavam legalmente obrigados à entrega da declaração de anual de ajuste (DIRPF). A grande maioria dos contribuintes com renda anual abaixo de R\$ 12.000,00 estava isenta pagamento do imposto. No exercício de 1999, quando os contribuintes do imposto de renda entregaram declaração referente aos rendimentos auferidos em 1998, um total aproximado de 11,5 milhões de declarações foram entregues. De acordo com o IBGE, a população economicamente ativa em 1998 era de aproximadamente 70 milhões de pessoas. Em razão desta diferença, os dados do histograma relativos à faixa de renda inferior a R\$ 12.000,00 foram excluídos da estimativa, por não corresponder ao universo de contribuintes com esta faixa de renda.

θ foi estimado em 1,68990, com erro padrão de 0.050 e R-quadrado de .99, apesar de rodado com 10 observações apenas.

Em 1998, os contribuintes com renda acima de R\$ 10.800,00, segundo os dados da Receita Federal, somavam 4.374.399 pessoas. Para uma renda tributável de R\$ 10.800,00, o valor de w correspondente é R\$ 13.285,07, calculado de acordo com as condições de primeira ordem resultantes de (16).

Obtido $E[Q_0(w)]$ a partir da fórmula (23), e presumindo que o governo necessita de uma quantia equivalente ao total arrecadado no ano com o imposto de renda, pode-se calcular uma alíquota marginal única que seja equivalente ao modelo de tributação sobre a renda existente. O valor calculado ficou em 14,97 % correspondendo à alíquota marginal única a ser aplicada sobre todos os contribuintes do imposto.

Para que se tenha uma compreensão da magnitude dos custos econômicos gerados pela tributação sobre a renda, utilizamos a simulação de um incremento marginal de 10 pontos percentuais sobre a maior alíquota de imposto de renda existente no ano calendário de 1998 (27,5%).

Com base nos dados disponíveis, considerando os contribuintes maximizando utilidade sobre a função descrita em (16), com parâmetro de aversão de risco de 1,1 e elasticidade de Frisch igual a 1, e considerando a presença tanto do efeito substituição quanto o efeito renda, a elasticidade compensada da renda tributável estaria próxima de 0,59.

A estes valores, uma receita tributária de R\$ 5,973 bilhões seria gerada a um custo adicional de R\$ 1,193 bilhões para a sociedade. Ou em outras palavras, a receita adicional de R\$ 5,836 bilhões tornaria a condição de vida dos contribuinte pior em R\$ 7.166 bilhões, o equivalente à receita tributária mais as perdas suportados em razão da ineficiência econômica. Cada real arrecadado pelo governo teria o custo extra de R\$ 0,20 para a sociedade, ou um custo total de R\$ 1,20 para que o governo arrecade R\$ 1,00.

Se o incremento marginal for o dobro, isto é, 20 pontos percentuais a mais na alíquota do imposto de renda, a perda de peso morto mais que dobra: uma receita tributária adicional de R\$ 8,965 bilhões seria gerada a um custo de eficiência de R\$ 5,367

bilhões. Cada real arrecadado teria um custo de R\$ 0,60 para a sociedade, ou um custo de R\$ 1,60 para cada R\$ 1,00 arrecadado pelo governo¹⁶.

Valores superiores de ETI podem fazer com que o real a mais arrecadado seja superado pelo custo estimado decorrente da perda de eficiência econômica introduzida pelo aumento de tributo.

¹⁶ O crescimento exponencial da perda de bem estar resultante do aumento sucessivo da alíquota marginal é previsto pela fórmula de Harberger (1964), de forma aproximada para pequenas variações de alíquota.

6 CONCLUSÃO

A obrigação de pagar tributos altera as decisões que as pessoas tomam em relação ao que consomem, poupam, o quanto investem em suas carreiras profissionais ou se lançarão mão de meios legais ou ilegais de contornar a exigência tributária. Esta afirmação é algo evidente e intuitiva, e qualquer indivíduo que paga imposto facilmente concordará com seu conteúdo.

Entretanto, grande parte das análises que amparam a adoção de políticas públicas que envolvam tributos frequentemente releva ou ignora esta condição elementar: contribuintes reagem quando são chamados a ceder recursos financeiros ao governo. Mesmo em países desenvolvidos a questão já foi subestimada; a título de exemplo, nos Estados Unidos apenas recentemente o Tesouro Americano e o Congresso passaram a considerar tais reflexos em suas análises e estimativas de receita que subsidiam a implantação de políticas tributárias.

A postura dos agentes públicos é compreensível. O objetivo imediato é a entrada de recursos para financiar os gastos governamentais. Frequentemente estimativas de receita são feitas assumindo a inexistência de efeitos contrários à tributação por parte dos contribuintes do imposto. Entretanto, além da carga tributária direta, os contribuintes suportam um custo adicional decorrente da tributação que não é diretamente observável. Por essa razão esse custo é geralmente negligenciado. As previsões “mecânicas” de arrecadação indicam que um aumento de alíquota marginal do imposto gerará um nível de receita, e a arrecadação efetiva é contabilizada nas contas nacionais. As possíveis perdas resultantes do comportamento reativo do contribuinte, e que podem ter dimensão considerável não são registradas contabilmente. Com um agravante: a carga tributária direta, arrecadada pelo governo, pode retornar na forma de bens e serviços públicos; o que não acontece com a carga indireta excedente e não contabilizada.

A grande dificuldade em obter estimativas de custo de eficiência resultante da imposição de um tributo é a variada gama de reações à tributação passíveis de serem adotadas pelo contribuinte. Conceber um modelo estrutural e estimar parâmetros

relacionados com todas estas reações pode ser bastante dificultoso. Neste sentido, a elasticidade da renda tributável é uma forma bastante simples de obter uma avaliação global de todos os efeitos conjugados, desde que se tenha conhecimento das limitações e restrições que envolvem este parâmetro. Chetty (2009) mostrou que relaxando algumas hipóteses, a ETI deixa de ser estatística suficiente para a avaliação de bem estar, que passa a depender da estimativa de novos parâmetros.

No presente trabalho, ao derivar um modelo de utilidade separável, levando em consideração a ocorrência do efeito renda, a ETI passa a variar de acordo com a política tributária a ser imposta pelo governo. Dependendo da política adotada pode-se falar de uma ETI antes da mudança, e uma ETI depois da mudança. Tal fato torna evidente que a ETI não é exógena ao sistema tributário vigente, e não pode ser caracterizada como um parâmetro estrutural.

Dificuldades à parte, a elasticidade da renda tributável oferece a possibilidade de estimar de forma relativamente simples medidas que não exijam alto grau de desagregação, como a perda de peso morto sofrida pela economia, ou previsões de arrecadação de tributos levando em conta aspectos comportamentais. No capítulo 5 exercitamos estes conceitos a partir da definição de uma função utilidade, e da utilização de dados agregados por faixa relativos ao imposto de renda sobre pessoas físicas do exercício de 1999, disponibilizados publicamente pela Receita Federal do Brasil. Foi observado que o aumento de alíquota marginal do imposto de renda produz um aumento mais que proporcional em termos de custo de eficiência,

Para que futuramente seja possível aprofundar pesquisas em extensão lógica a este trabalho, de fundamental importância é a utilização da base de dados da Receita Federal para a montagem de conjuntos de dados de corte transversal e dados de painel a partir de informações dos contribuintes. Estes dados permitiriam a aplicação de métodos empíricos mais sofisticados e mais precisos, visando a investigação e avaliação dos componentes das reações comportamentais à tributação, e a sua influência na soma geral de fatores comportamentais que impactam a arrecadação de tributos no curto e longo prazos. Outras nações, desenvolvidas ou em desenvolvimento movimentam-se no sentido de aperfeiçoar suas estruturas tributárias, questão prioritária

tendo em vista o ambiente de competição econômica acirrada entre países presente nos dias de hoje, e possivelmente cada vez mais intensa em dias futuros. É grande a importância de se adicionar este conhecimento ao arsenal de informações necessárias à elaboração de políticas públicas que visem um sistema tributário mais eficiente e equitativo, que efetivamente contribua na criação de condições para o desenvolvimento socio-econômico consistente do país.

REFERENCIAS BIBLIOGRAFICAS

AUERBACH, Alan J. The Theory of Excess Burden And Optimal Taxation. Handbook of Public Economics, V 1 A. Auerbach e M Feldstein (eds.), North Holland Publishers, 1985.

AUTEN, Gerald; CARROLL, Robert. The Effect of Income Taxes on Household Behavior. Review of Economics and Statistics, 81(4): 681-693,1999.

ATKINSON, Anthony B.; STIGLITZ, Joseph E. Lectures on Public Economics; MacGraw-Hill, 1980.

CHETTY, Raj. A New Method of Estimating Risk Aversion. American Economic Review 96(5), 1821-1834, 2006

CHETTY, Raj. Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," American Economic Journal: Economic Policy 1(2): 31-52, 2009.

CHETTY, Raj. Sufficient Statistics For Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. Annual Review of Economics 1: 451-488, 2009b

COSTA, Carlos Eugênio da; PEREIRA, Thiago Neves. The Optimal Labor Income Tax Schedule For Brazil, Working Paper, 2010.

CREEDY, John. Measuring Welfare Changes And The Excess Burden of Taxation. Bulletin of Economic Research 52:1, 2000.

DAHLBY, Bev. Progressive Taxation And The Social Marginal Cost Of Public Funds. Journal of Public Economics, 67: 105-122, 1998.

FELDSTEIN, Martin. The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. Journal of Political Economy, 103(3): 551-572, 1995.

FELDSTEIN, Martin. Tax Avoidance and the Deadweight Loss of the Income Tax. Review of Economics and Statistics, 81(4): 674-680, 1999.

GOOLSBEE, Austan. Evidence On The High-Income Laffer Curve From Six Decades Of Tax Reform. *Brooking Papers On Economic Activity* 2: 1-47, 1999.

GOOLSBEE, Austan. What Happens When You Tax the Rich? Evidence from Executive Compensation. *Journal of Political Economy*, 108(2): 352-378, 2000.

GORODNICHENKO, Yuriy; MARTINEZ-VAZQUEZ, Jorge; SABIRIANOVA, Klara. Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia, *forthcoming Journal of Political Economy*, 2009.

GRUBER, Jonathan; SAEZ, Emmanuel. The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics*, 84(1): 1-32, 2002.

HARBERGER, Arnold. Taxation, Resources Allocation, and Welfare; J Due (ed) *The Roles of Direct and Indirect Taxes in the Federal Tax System*, Princeton University Press, 1964.

HAUSMAN Jerry A. The Econometrics of Nonlinear Budget Sets; *Econometrica* Vol. 53, No. 6 pp. 1255-1282, 1985.

HAUSMAN, Jerry A. Taxes and Labor Supply. *Handbook of Public Economics*, V 1 A. Auerback e M Feldstein (eds.), North Holland Publishers, 1985.

IMMERVOLL, Herwig; LEVY, Horácio; NOGUEIRA, José Ricardo; O'DONOGHUE, Cathal; SIQUEIRA, Rozane. Simulating The Impact of Inflation on The Prograssivity of Personal Income Tax in Brazil. *Revista Brasileira de Economia*, vol 64, No. 4 pp. 405-422 2010.

KIMBALL, Miles S.; SHAPIRO, Matthew D. Labor Supply: Are The Income And Substitution Effects Both Large Or Both Small? Working Paper, National Bureau Of Economic Research, 2008.

KOPCZUK, Wojciech. Tax Bases, Tax Rates and the Elasticity of Reported Income. *Journal of Public Economics*, 89(11-12): 2093-2119, 2005.

MEGHIR, Costas; PHILLIPS, David; Labour Supply and Taxes. *Dimensions of Tax*

Design – The Mirrlees Review, Chaired by Sir James Mirrlees for Institute for Fiscal Studies; Oxford University Press, 2010.

SAEZ, Emmanuel. Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies*, 68: 205-229, 2001

SAEZ, Emmanuel. The Effect of Marginal Tax Rates on Income: A Panel Study of “Bracket Creep”. *Journal of Public Economics*, 87: 1231-1258, 2003.

SAEZ, Emmanuel. Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications. *Tax Policy and the Economy*, Vol. 18, ed. James Poterba. Cambridge: MIT Press, 2004.

SAEZ, Emmanuel; SLEMROD, Joel; GIERTZ, Seth. The elasticity of taxable income with respect to marginal tax rates: a critical review. *National Bureau of Economic Research Working Paper*, no. 15012, 2009.

SLEMROD, Joel; YITZHAKI, Shlomo. Tax Avoidance, Evasion and Administration. in Alan Auerbach and Martin Feldstein, ed., *Handbook of Public Economics*, v3, North-Holland: 1423-1470, 2002.

SLEMROD, Joel; KOPCZUK, Wojciech. The Optimal Elasticity of Taxable Income. *Journal of Public Economics*, 84(1): 91-112, 2002.

SECRETARIA DA RECEITA FEDERAL. Análise Econômica da DIRPF 1999, Texto para Discussão 09 - disponível em www.receita.fazenda.gov.br, 2000.

SECRETARIA DA RECEITA FEDERAL. Considerações sobre o imposto de renda da pessoa física no Brasil - disponível em www.receita.fazenda.gov.br, 2001.

SECRETARIA DA RECEITA FEDERAL. O Imposto de Renda das Pessoas Físicas no Brasil, Estudos Tributários 14 – disponível em www.receita.fazenda.gov.br, 2004.

SECRETARIA DA RECEITA FEDERAL. Memória Receita Federal - Imposto de Renda Pessoa Física - disponível em www.receita.fazenda.gov.br/Memoria/irpf/default.asp.

SECRETARIA DA RECEITA FEDERAL. Carga Tributária no Brasil – de 1994 a 2010 - disponível em www.receita.fazenda.gov.br.

ZEE, Howell H. Reforma do Imposto de Renda Pessoa Física: Conceitos, Desafios e Evolução Comparativa em Diversos Países, IMF Working Paper, 2005.

APÊNDICE A

Genericamente, a variação equivalente assume a forma da seguinte igualdade de funções despesa:

$$e(p_0, v(p_1, m)) = e(p_0, v(p_0, m - EV))$$

onde p_0 e p_1 são vetores de preços iniciais e finais, m a renda total do indivíduo, $v(p_1, m)$ a função utilidade indireta, e EV a variação equivalente. Considerando as relações

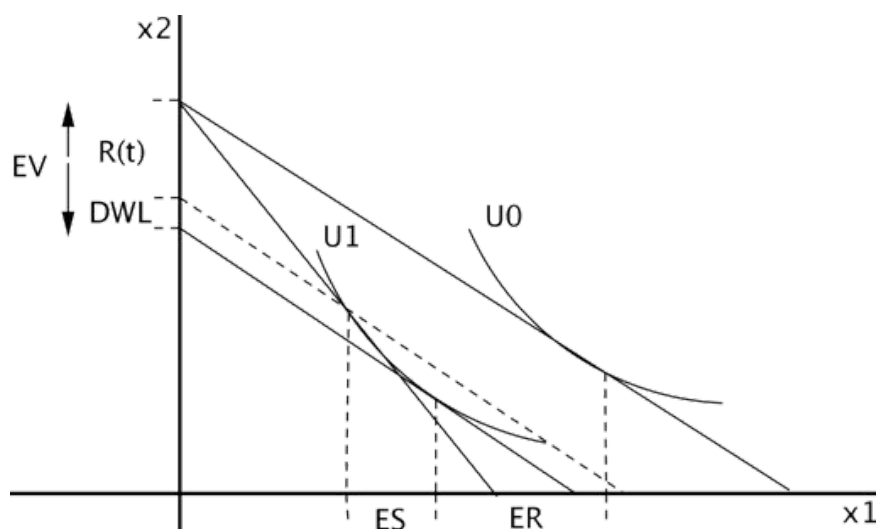
$$e(p_0, v(p_0, m - EV)) = m - EV$$

$$e(p_1, v(p_1, m)) = m$$

podemos escrever a variação equivalente como

$$EV = e(p_1, v(p_1, m)) - e(p_0, v(p_1, m))$$

A figura a seguir ilustra, através da tributação de uma commodity x_1 , como os conceitos de perda de peso morto, variação equivalente, receita tributária, tributação lump sum e efeitos renda e substituição estão estreitamente relacionados:



A figura mostra claramente que a totalidade da perda de peso morto (DWL) pode ser atribuída exclusivamente ao efeito substituição. O resultado da variação equivalente (EV) é igual a imposição de uma tributação lump sum que leva o indivíduo de U_0 a U_1 . A receita efetivamente arrecadada pelo governo corresponde a $R(t)$.

O efeito substituição é exclusivamente responsável pela perda de peso morto, e é gerado pelo deslocamento do ponto tangente e a mudança da inclinação da restrição orçamentária ao longo da curva de indiferença.

APÊNDICE B

Considerando t um tributo a uma alíquota fixa, que incide de forma proporcional sobre a renda do trabalho, o problema do consumidor é:

$$\begin{aligned} \max_{C,L} U(C,L) \\ \text{s.a} \\ C = Lw(1-t) + I \end{aligned}$$

Do problema de maximização da utilidade, a partir das condições de primeira ordem

$$\begin{aligned} U'_C(C,L) &= \lambda \\ U'_L(C,L) &= -w_0\lambda \\ w_0L + y - C &= 0 \end{aligned}$$

obtemos as funções marshallianas $L = L^*(w,I)$ e $C = C^*(w,I)$. Analogamente, pelo problema de minimização de I ,

$$\begin{aligned} \min_{C,L} I = C - w_0L \\ \text{s.a} \\ U(C,L) \geq \bar{U} \end{aligned}$$

extraímos o lagrangeano:

$$\ell = C - w_0L + \mu(U_0 - U(C,L))$$

e as funções hicksianas $L = L^H(w_0, U)$ e $C = C^H(w_0, U)$.

Através da identidade $L^H(w_0, U) = L^*(w_0, I(w_0, U))$, e diferenciando em relação a w_0 :

$$\frac{\partial L^H}{\partial w_0} = \frac{\partial L^*}{\partial w_0} + \frac{\partial L^*}{\partial I} \frac{\partial I}{\partial w_0}$$

do lagrangeano, utilizando o teorema do envelope:

$$\frac{\partial y}{\partial w_0} = \frac{\partial \ell}{\partial w_0} = -L$$

A equação de Slutsky escrita em termos de oferta do trabalho é

$$\left. \frac{\partial L}{\partial w_0} \right|_I = \left. \frac{\partial L}{\partial w_0} \right|_U + L \frac{\partial L}{\partial I}$$

O primeiro termo do lado direito da equação é o termo compensado, responsável pelo efeito substituição; o segundo termo é o efeito renda, ponderado pela quantidade de trabalho (em horas) ofertada.

APÊNDICE C

Neste apêndice, a partir da maximização de uma função utilidade separável entre o consumo C e a renda Q obtemos fórmulas específicas para a elasticidade da renda tributável, elasticidade-renda em relação a w , e o parâmetro η do efeito renda, utilizado na equação de Slutsky (5).

Começamos do problema apresentado em (16)

$$\max U(C, Q) = \frac{C^{1-\rho}}{1-\rho} - \frac{1}{\gamma} \left(\frac{Q}{w} \right)^\gamma,$$

$$\text{sujeito a } C = (1-t)Q + I,$$

A condição de primeira ordem:

$$[(1-t)Q + I]^{-\rho}(1-t) - \frac{Q^{\gamma-1}}{w^\gamma} = 0, \text{ ou}$$

$$-\rho \ln[(1-t)Q + I] + \ln(1-t) - (\gamma-1) \ln Q + \gamma \ln w = 0$$

Diferenciando,

$$\frac{\partial f(Q, I, w, (1-t))}{\partial(1-t)} \frac{d(1-t)}{d(1-t)} + \frac{\partial f(Q, I, w, (1-t))}{\partial Q} \frac{dQ}{d(1-t)} + \frac{\partial f(Q, I, w, (1-t))}{\partial I} \frac{dI}{d(1-t)} + \frac{\partial f(Q, I, w, (1-t))}{\partial w} \frac{dw}{d(1-t)} = 0$$

assumindo que I e w são independentes de $(1-t)$:

$$\frac{\partial f(Q, I, w, (1-t))}{\partial(1-t)} + \frac{\partial f(Q, I, w, (1-t))}{\partial Q} \frac{dQ}{d(1-t)} = 0$$

$$\frac{dQ}{d(1-t)} = - \frac{\partial f(Q, I, w, (1-t)) / \partial(1-t)}{\partial f(Q, I, w, (1-t)) / \partial Q} \quad (I)$$

da mesma forma:

$$\frac{dQ}{dI} = - \frac{\partial f(Q, I, w, (1-t)) / \partial I}{\partial f(Q, I, w, (1-t)) / \partial Q} \quad (II)$$

$$\frac{dQ}{dw} = - \frac{\partial f(Q, I, w, (1-t))/\partial w}{\partial f(Q, I, w, (1-t))/\partial Q} \quad (\text{III})$$

derivando a função em relação a cada variável

$$\frac{\partial f(Q, I, w, (1-t))}{\partial (1-t)} = \frac{-\rho Q}{Q(1-t) + I} + \frac{1}{(1-t)}$$

$$\frac{\partial f(Q, I, w, (1-t))}{\partial Q} = \frac{-\rho(1-t)}{Q(1-t) + I} + \frac{(\gamma-1)}{Q}$$

$$\frac{\partial f(Q, I, w, (1-t))}{\partial I} = \frac{-\rho}{Q(1-t) + I}$$

$$\frac{\partial f(Q, I, w, (1-t))}{\partial w} = \frac{\gamma}{w}$$

Com o emprego das funções derivadas acima e das diferenciações I, II e III, obtemos as seguintes fórmulas:

- Elasticidade da renda tributável:

$$-\frac{\partial f(.)/\partial (1-t)}{\partial f(.)/\partial Q} = -\frac{\frac{-\rho Q}{Q(1-t) + I} + \frac{1}{(1-t)}}{\frac{-\rho(1-t)}{Q(1-t) + I} + \frac{(\gamma-1)}{Q}} = \frac{dQ}{d(1-t)}$$

$$\frac{dQ}{d(1-t)} = \frac{\frac{\rho Q(1-t) - Q(1-t) + I}{(Q(1-t) + I)(1-t)}}{\frac{\rho(1-t)Q + (\gamma-1)(Q(1-t) + I)}{(Q(1-t) + I)Q}} =$$

$$\frac{dQ}{d(1-t)} \frac{(1-t)}{Q} = \frac{\rho Q(1-t) - Q(1-t) + I}{\rho(1-t)Q + (\gamma-1)(Q(1-t) + I)} = \xi$$

- Elasticidade da renda em relação a w:

$$\frac{dQ}{dw} = - \frac{\partial f(.) / \partial w}{\partial f(.) / \partial Q} = - \frac{\frac{\gamma}{w}}{\frac{-\rho(1-t)}{Q(1-t)+I} - \frac{(\gamma-1)}{Q}} =$$

$$\text{tomando } \hat{\rho} = \frac{\rho(1-t)Q}{Q(1-t)+I}$$

$$\frac{dQ}{dw} = \frac{\gamma/w}{(\hat{\rho} + \gamma - 1)/Q}$$

$$\frac{dQ}{dw} \frac{w}{Q} = \frac{\gamma}{(\hat{\rho} + \gamma - 1)} \quad (21)$$

- Efeito renda

$$\frac{dQ}{dI} = - \frac{\partial f(.) / \partial I}{\partial f(.) / \partial Q} = - \frac{\frac{-\rho}{Q(1-t)+I}}{\frac{-\rho(1-t)}{Q(1-t)+I} - \frac{(\gamma-1)}{Q}} =$$

usando $\hat{\rho}$

$$\frac{dQ}{dI} = - \frac{\hat{\rho}/Q(1-t)}{(\hat{\rho} + \gamma - 1)/Q}$$

$$\frac{dQ}{dI} (1-t) = - \frac{\hat{\rho}}{(\hat{\rho} + \gamma - 1)} = \eta$$

a elasticidade compensada da renda tributável pode, consequentemente, ser encontrada através da equação de Slutsky:

$$\left. \frac{dQ}{d(1-t)} \frac{(1-t)}{Q} \right|_U = \frac{dQ}{d(1-t)} \frac{(1-t)}{Q} - \frac{dQ}{dI} (1-t) =$$

$$\xi^h = \xi^m - \eta = \frac{\rho Q(1-t) - Q(1-t) + I}{\rho(1-t)Q + (\gamma-1)(Q(1-t)+I)} + \frac{\hat{\rho}}{(\hat{\rho} + \gamma - 1)}$$

APÊNDICE D

O rendimento dos indivíduos antes da imposição do tributo sobre a renda Q_0 , é função da habilidade ou qualificação do indivíduo (w). Para modelar a distribuição de w , foi utilizada a Distribuição de Pareto. O objetivo é encontrar uma formulação que defina o valor esperado de Q_0 . A distribuição de Pareto acumulada é

$$F(w) = 1 - \left(\frac{k}{w} \right)^\theta$$

e a função densidade de distribuição é $f(w) = \frac{\theta k^\theta}{w^{\theta+1}}$.

Se $k \leq w \leq \infty$, e $\theta > \hat{\gamma}$, $\gamma/(\rho + \gamma - 1) = \hat{\gamma}$

$$\int_k^\infty w^{\gamma/(\rho+\gamma-1)} \frac{\theta k^\theta}{w^{\theta+1}} dw =$$

$$\theta k^\theta \int_k^\infty w^{\left[\frac{\gamma}{(\rho+\gamma-1)} - (\theta+1) \right]} dw =$$

$$\theta k^\theta \left(\frac{w^{\left[\frac{\gamma}{(\rho+\gamma-1)} - \theta \right]}}{\left[\frac{\gamma}{(\rho+\gamma-1)} - \theta \right]} \right) \Bigg|_k^\infty =$$

$$\theta k^\theta \left(\frac{k^{\left[\frac{\gamma}{(\rho+\gamma-1)} - \theta \right]}}{\left[\theta - \frac{\gamma}{(\rho+\gamma-1)} \right]} \right) =$$

$$\frac{\theta(\rho + \gamma - 1)k^{\left[\frac{\gamma}{(\rho+\gamma-1)} \right]}}{\left[\theta(\rho + \gamma - 1) - \gamma \right]} = E[Q_0(w)]$$