

A EVOLUÇÃO DOS FUNDAMENTOS TEÓRICOS DA POLÍTICA MONETÁRIA NO ENFOQUE MONETARISTA*

Aod Cunha de Moraes Jr.**

O escopo principal deste trabalho é apresentar a evolução dos fundamentos teóricos que justificam a opção por controles quantitativos do estoque de moeda, dentro de uma linha do pensamento macroeconômico que postula para a política monetária a tarefa do controle inflacionário e a rejeição a estímulos para o crescimento do produto e do emprego.

Dentro dessa proposta, serão agrupados três conjuntos de proposições pertencentes, respectivamente, às formulações da Teoria Quantitativa da Moeda (TQM) no início do século XX, à obra de Friedman nas décadas de 50 e 60 e às contribuições da Economia Novo-Clássica. Deve-se salientar que não há interesse neste artigo de fazer uma avaliação crítica do arcabouço teórico em questão, mas, sim, apresentá-lo, em cada uma de suas etapas, com uma clareza que permita observarem-se as hipóteses e as conclusões que traçam novos contornos ao espaço definido para a política monetária. Será explorada uma trajetória que inclui desde a construção de uma teoria da demanda por moeda até um modelo mais completo de determinação da renda, com suposições mais sofisticadas sobre o comportamento da oferta agregada.

1 - A Teoria Quantitativa da Moeda no início do século XX

A idéia de que a quantidade de moeda em circulação afeta o nível geral de preços e de que há uma dicotomização dos mercados monetário e real não se inicia neste

* Este texto corresponde a uma etapa de uma pesquisa mais ampla proposta no artigo **O Controle da Oferta de Moeda: Monetaristas Keynesianos Neoclássicos e Pós-Keynesianos**, que foi apresentado pelo autor no XXI Encontro Nacional dos Cursos de Pós-Graduação em Economia, realizado em Belo Horizonte.

** Economista da Secretaria da Fazenda e Professor da PUC-RS.

O autor agradece aos professores Roberto C. Moraes, Pedro Cezar D. Fonseca e Fernando Ferrari, do CPGE-UFRGS, por comentários feitos numa fase inicial da pesquisa, e, em especial, ao colega André Minella, cujas sugestões e incentivos foram decisivos para uma versão final deste artigo. Qualquer equívoco remanescente é de inteira responsabilidade do autor.

século.¹ No entanto a grande maioria das referências feitas na Macroeconomia à Teoria Quantitativa da Moeda relaciona-se com as versões mais elaboradas do início do século, onde se destacam a equação de trocas de Fisher (1911) e a equação de Cambridge — fruto dos trabalhos de Marshall e Pigou, entre outros.

Tanto a equação de trocas como a de Cambridge, que serão apresentadas a seguir, sugerem ilações importantes sobre o controle da oferta de moeda.

1.1 - A equação de trocas

Na formulação de Fisher (1911), o papel da moeda como meio de troca não chega a ser explicitado em termos de uma teoria de demanda por moeda, e o conceito velocidade de circulação da moeda aparece na versão "transacional".

$$M_s V_t = PT \quad (1)$$

onde

M_s = quantidade de moeda;

V_t = velocidade de circulação da moeda expressa no número de transações realizadas por um dado estoque de moeda num dado período de tempo;

T = volume de transações;

P = nível de preços.

A medida que a equação (1) expressa uma identidade contábil onde, em um dado período, o valor do fluxo de compras é igual ao valor das vendas realizadas, passa a ser teoricamente relevante saber o que determina cada um dos termos da equação. De acordo com Fisher, a quantidade de moeda M_s é determinada independentemente de qualquer uma das outras três variáveis e a qualquer momento pode ser considerada como dada. Supondo-se que o único nível de renda de equilíbrio é o de pleno emprego — vale a Lei de Say —, assume-se que há uma razão estável entre o volume de transações T , também considerado como dado, e o nível de produto. A velocidade de circulação da moeda V_t é tratada como sendo independente das outras três variáveis e, ainda que não seja necessariamente constante ao longo do tempo, convergirá rapidamente para o seu valor de equilíbrio,² quando houver qualquer distúrbio inicial.

¹ Ver Moraes (1991) para uma abordagem mais detalhada.

² Esse valor é determinado por fatores institucionais, como o nível de organização do sistema bancário.

Feitas as suposições de que T e o valor de equilíbrio de V_t são constantes, chega-se à Teoria Quantitativa da Moeda, onde o nível de preços de equilíbrio é determinado de forma proporcional pela quantidade de moeda.

$$M_s \bar{V}_t = P\bar{T} \quad (2)$$

Seguindo-se Laidler (1978), pode se acrescentar que, mesmo não tendo sido proposto por Fisher, a equação (2) sugere um comportamento da demanda por moeda do tipo

$$M_d = k_t P\bar{T} \quad (3)$$

que, somada à condição de equilíbrio no mercado monetário,

$$M_d = M_s \quad (4)$$

leva a

$$M_s \frac{1}{k_t} = M_s \bar{V}_t = P\bar{T} \quad (5)$$

onde

$$\bar{V}_t = \frac{1}{k_t} \quad (6)$$

Dessa forma, há uma função de demanda por moeda que relaciona a retenção de saldos monetários para fins transacionais. Se \bar{T} e \bar{V}_t permanecem constantes, então variações em P serão acompanhadas de variações proporcionais em M_d , o que equivale dizer que a demanda real por moeda (M_d/P) também permanece constante.

1.2 - A equação de Cambridge

Nessa versão da TQM, há uma postulação de ordem microeconômica para a demanda por moeda, ao substituir-se o "enfoque transacional" para um "enfoque de encaixe":

"The Cambridge economists did not ask, as Fisher did, what determines the amount of money an economy needs to carry out a given volume of transactions, but rather what determines the amount of money an individual would wish to hold given that the desire to conduct transactions makes money holding desirable at all" (LAIDLER, 1978, p.59).

[*"Os economistas de Cambridge não perguntaram, como Fisher fez, o que determina a quantidade de moeda que uma economia necessita para financiar um determinado volume de transações, mas principalmente o que determina a*

quantidade de moeda que um indivíduo deseja manter, dado que o desejo de realizar transações faz com que a posse da moeda seja atraente.]"

A equação da demanda por moeda aparece como

$$M_d = kPY \quad (7)$$

onde Y = renda real agregada, acrescentando a condição de equilíbrio no mercado monetário,

$$M_d = M_s \quad (8)$$

resulta que

$$M_s = kPY \quad (9)$$

portanto,

$$M_s \frac{1}{k} = M_s V = PY \quad (10)$$

onde, em comparação com a versão de Fisher, ao invés de V representar a **velocidade-transação** da circulação da moeda passa a representar a **velocidade-renda** da circulação da moeda ou, simplesmente, a velocidade-renda — não o número de vezes que a unidade física da moeda se movimenta, mas a proporção entre a taxa de circulação do estoque monetário e a taxa de produção em termos da renda real.

Na equação (7), o termo k é a "razão desejada" entre a liquidez real demandada (M_d/P) e a renda real Y . Nessa formulação, a velocidade-renda da moeda, interpretada como o inverso de k , apresenta uma clara vantagem operacional do ponto de vista estatístico, visto que escapa da difícil mensuração do volume de transações. Permanece, no entanto, válida a mensagem central da TQM de que as oscilações no estoque monetário se comunicam exclusivamente, de forma diretamente proporcional, com as oscilações no nível de preços.

1.3 - Oscilações na oferta de moeda

Tanto na equação de Fisher como, principalmente, na equação de Cambridge, são possíveis considerações sobre algumas conseqüências da oscilação da oferta monetária.

Como se viu, a suposição de que V e T e a de que k e Y sejam constantes no curto prazo são fundamentais à proposição quantitativa. Dentro dessa suposição, há ainda espaço para três casos:

- a) pode-se compreender "constante" como a manutenção de um equilíbrio estático para todas as variáveis;

- b) pode-se compreender "constante" como a existência de um equilíbrio dinâmico em torno de um valor no curto prazo; e
- c) para os casos (a) e (b), pode existir um padrão diferenciado de comportamento entre as variáveis.

Analisando-se a equação de Cambridge, no caso (a), o aumento ou a diminuição da oferta monetária repercutiria instantaneamente apenas no nível de preços. No caso (b), até que todas as variáveis se ajustassem a uma mudança no estoque monetário, poderiam haver efeitos transitórios em k e Y . De acordo com (c), poderia ocorrer uma maior sensibilidade de k ou Y ou, ainda, de apenas uma das duas. Assim, a forma com que as variações da oferta de moeda se traduzem no curto prazo, em variações no nível de preços, depende do comportamento dinâmico de k e Y em torno de seus valores de equilíbrio. Note-se, por exemplo, que a proposição de neutralidade da oferta monetária em relação à renda real agregada pode persistir mesmo quando k não permanece fixo, ao mesmo tempo em que a estabilidade do equilíbrio de k não garante a neutralidade da moeda.

Pode ser chamada atenção para o fato de que a teoria quantitativa, na formalização ora analisada, não apresenta como elemento inovador o axioma de que a moeda é neutra — este já está "assegurado" pelos postulados da economia clássica, que garantem o pleno emprego na estrutura de produção da economia. Mas, ao se supor que a moeda, não rendendo juros e sendo indesejável enquanto aplicação patrimonial, é demandada apenas pelo seu atributo facilitador das trocas, a TQM expressou a hipótese de uma demanda real por moeda constante no curto prazo. Assim, todo o aumento do estoque monetário terá que ser seguido, na mesma proporção, por um aumento no nível de preços, e toda a diminuição do estoque monetário, por uma diminuição no nível de preços. A estabilidade do nível de preços requer que a oferta monetária acompanhe, na mesma proporção, o crescimento real do produto, só sendo recomendável algum crescimento maior ou menor da oferta de moeda se houver alguma mudança significativa no arranjo institucional que determina k .

2 - Friedman: a reconstrução da TQM e a quantidade ótima de moeda

Após a crise de 1929 e a Depressão dos anos 30, o desenvolvimento da teoria keynesiana abalou a hegemonia da ortodoxia clássica e, dentro dela, a própria ortodoxia monetária representada pela TQM. Abriu-se espaço para importantes prescrições de política econômica na linha da sintonia fina do pleno emprego, conhecida como *demand management approach*.³

³ Termo que se utiliza na literatura para a recomendação do uso de instrumentos fiscais e monetários no controle da demanda agregada.

Não só os postulados da economia clássica que garantiam o pleno emprego foram minados a partir das críticas de Keynes (1936) em seu Capítulo 2, mas junto com o arcabouço analítico do princípio da Demanda Efetiva apareciam outras inovações teóricas. Dentre elas, um novo formato para o comportamento da demanda por moeda sugeria que esta seria sensível à taxa de juros e que esse seria um caminho explorável do ponto de vista da política monetária — ainda que a associação constante com o elevado nível de desemprego da década de 30 conduza ao caso possível da armadilha da liquidez e, conseqüentemente, à primazia da política fiscal. Outra sugestão importante para a teoria monetária era a de que, sendo os preços rígidos, ou, pelo menos, sem a total flexibilidade walrasiana, um aumento na oferta de moeda, dado um estado de expectativas quanto ao futuro, ou seja, dada uma curva de preferência pela liquidez, poderia levar a um crescimento da demanda real por moeda. O controle da oferta de moeda poderia estar subordinado à busca de efeitos reais positivos no nível de emprego e de renda através do monitoramento das taxas de juros.

Foi em meio ao sucesso intelectual e político das recomendações keynesianas que emerge uma nova versão para a Teoria Quantitativa da Moeda, postulando, nas palavras do seu mais famoso defensor, recuperar o espaço perdido:

"The present volume is partly a symptom of this reemergence and partly a continuance of an aberrant tradition. Chicago was one of the few academic centers at which the quantity theory continued to study monetary problems. The quantity theory that retained this role differed sharply from the atrophied and rigid caricature that is so frequently described by the proponents of the new income — expenditure approach — and with some justice, to judge by much of the literature on policy that was spawned by quantity theorists" (FRIEDMAN, 1956, p.3).

["O presente volume é em parte um sintoma desta reemergência e em parte a continuação de uma forte tradição. Chicago foi um dos poucos centros acadêmicos onde a teoria quantitativa continuou a estudar problemas monetários. A teoria quantitativa que conservou este papel difere nitidamente da atrofiada e rígida caricatura que é tão freqüentemente descrita pelos proponentes da nova versão "renda-gasto" — e com alguma justiça, considerada por boa parte da literatura que foi difundida por teóricos quantitativistas."]

A reconstrução friedmaniana da TQM, com suas decorrentes implicações para a política monetária, é encontrada basicamente em dois textos: **The quantity theory of money — A restatement** (1956) e **The demand for money: some theoretical and empirical results** (1959). No primeiro, Friedman trata de construir uma argumentação teórica para uma função de demanda por moeda, onde a sua estabilidade, antes de ser representada por um valor constante para a velocidade-renda de circulação da moeda, deve ser entendida a partir da estabilidade da função que determina o valor da velocidade-renda de circulação da moeda, ou seja, que esta possa ser razoavelmente explicada tomando-se como base apenas as variáveis explicitadas na função. No

segundo artigo, Friedman parte para o teste empírico de sua função, ao mesmo tempo em que pretende demonstrar que a introdução da noção de "renda permanente", ao invés de "renda corrente", explica a aparente contradição entre o movimento secular que relaciona o estoque de moeda e a renda nominal e o movimento ao longo dos ciclos observado por Seldens (1956). A partir das constatações do artigo de 1959, Friedman retira conclusões importantes para a teoria da moeda, para o ciclo dos negócios e para a conduta da política monetária.

Como o próprio Friedman já havia antecipado no seu artigo de 1956, para que sua nova versão da TQM — uma teoria da demanda por moeda, como ele a definiu — pudesse chegar a um modelo completo de determinação da renda, seria preciso completá-lo com suposições adicionais sobre a estrutura da oferta agregada. Essa tarefa, junto com a prescrição de uma quantidade ótima de moeda, vem com a incorporação da análise da Curva de Phillips ampliada pelas expectativas adaptativas, levando-o às conclusões apresentadas em *The Role of Monetary Policy* (FRIEDMAN, 1968) e em *The Theoretical Framework for Monetary Analysis* (FRIEDMAN, 1970).

2.1 - A estabilidade da demanda por moeda

Friedman (1956) trata a demanda por moeda da mesma forma que trataria a demanda por qualquer bem durável, onde a utilidade é compreendida pelo fluxo de serviços que rende ao seu possuidor. Como qualquer outro caso de aplicação da teoria da demanda, deveria ser analisada a natureza da restrição orçamentária e selecionadas as variáveis relevantes para a mensuração do custo de oportunidade em se reter moeda. Assim, a demanda por moeda dependeria de três conjuntos de fatores: (a) a riqueza total mantida nas suas várias formas — analogamente à restrição orçamentária; (b) o preço e o retorno dessa forma específica e de formas alternativas; (c) o gasto e a preferência dos possuidores finais de riqueza.

Da especificação dos diferentes fluxos de renda gerados pelas formas alternativas de retenção de riqueza, chega-se à seguinte função de demanda por moeda:

$$M = f\left(P, i_n - \frac{1}{i_n} \frac{di_n}{dt}, i_r + \frac{1}{P} \frac{dP}{dt} - \frac{1}{i_c} \frac{di_c}{dt}, \frac{1}{P} \frac{dP}{dt}, W, \frac{Y}{i}, U\right) \quad (11)$$

onde

P = nível de preços;

i_n = taxa de juros dos títulos de renda nominal constante;

i_r = taxa de juros dos títulos de renda real constante;

W = relação entre os rendimentos associados ao capital não humano e a queles associados ao capital humano;

i_e = taxa de juros "média" da economia;⁴
 Y = fluxo total de rendas em termos nominais;
 U = variável que capta gostos e preferências;

ou, após algumas transformações algébricas, chega-se a

$$Y = v (i_n, i_r, \frac{1}{P} \frac{dP}{dt}, W, \frac{Y}{P}, U)M \quad (12)$$

onde a equação fica expressa na forma usual da teoria quantitativa e v é a velocidade-renda.

Friedman faz questão de ressaltar que, apesar de não ter sido explicitado nada sobre o comportamento dos bancos ou produtores de dinheiro, isso não afetaria sua análise, visto que repercutiria sobre a quantidade de moeda ofertada. Dada essa quantidade, a equação (1-2) define as condições sob as quais esse estoque de moeda será demandado.

No artigo de 1959, Friedman procura demonstrar a estabilidade da sua função de demanda por moeda, utilizando a noção de "renda permanente", quando explica que o comportamento da velocidade-renda no trabalho de Seldens (1956) é perfeitamente compatível com essa suposição. De acordo com a hipótese de renda permanente, quando um consumidor experimenta um aumento transitório de sua renda, isto é, quando sua renda observada em um período exceder o fluxo médio passado de sua renda, a parte transitória é acrescida aos seus ativos ou usada para reduzir seus passivos, sem alterar o padrão de consumo. Surge, também, uma importante separação na determinação do estoque nominal de moeda do estoque real. Os detentores de moeda não podem alterar o estoque nominal de moeda criado pela autoridade monetária, mas podem determinar o estoque real que desejam, na medida em que gastam mais e elevam a renda monetária e os preços.

Utilizando-se de dados do período de 1870 a 1954, Friedman estimou duas equações básicas:

$$\frac{M}{NP_p} = V \left(\frac{Y_p}{NP_p} \right)^{\gamma} \quad (13)$$

⁴ Segundo Friedman (1956, p.9), "This general rate, r , is to be interpreted as something of a weighted average of the two special rates plus the rates applicable to human wealth and to physical goods. Since the latter two cannot be observed directly, it is perhaps best to regard them as varying in some way with $r_b \dots$ and r_e ".

["(...) esta taxa geral, r , é para ser interpretada como algo próximo à média ponderada das duas taxas especiais mais as taxas aplicáveis à riqueza humana e aos bens físicos. Dado que as últimas duas não podem ser observadas de forma direta, é provavelmente melhor considerá-las variando de alguma forma com r_b e r_e .]

e

$$V = \frac{Y}{Y_p} \frac{1}{V} \left(\frac{Y_p}{N} \right)^{1-\gamma} \quad (14)$$

sendo que

$$Y_p(T) = \beta \int_{-\infty}^T e^{(\beta-\infty)(t-T)} Y(t) dt \quad (15)$$

onde

N = população;

Y = renda observada medida em termos nominais;

P = nível observado de preços;

M = estoque de moeda em termos nominais, onde os valores observados e permanentes são tratados como idênticos;

Y_p, P_p = renda agregada permanente em termos nominais e o nível de preços permanente respectivamente;

V = Y/M = velocidade observada;

V_p = Y_p/M = velocidade permanente.

A equação (15) estima a renda esperada no período T através de uma média ponderada dos valores passados da renda, ajustada pelo crescimento secular em termos de percentual anual, onde os pesos declinam exponencialmente. Os resultados obtidos por Friedman levaram-no a concluir que os movimentos cíclicos da velocidade-renda refletem, com segurança, movimentos ao longo de uma curva de demanda por moeda estável, sendo que as aparentes discrepâncias entre o movimento secular e os resultados para o ciclo refletem divergências entre os valores observados da renda e dos preços que servem às construções estatísticas e aquelas magnitudes que servem para que os detentores de moeda ajustem seus encaixes.

Dentre as conclusões retiradas do artigo de 1959, critica-se a ênfase dada à diferenciação dos motivos da demanda por moeda: motivo transação, motivo especulação e motivo precaução. Friedman não vê razão para essa divisão, à medida que julga a moeda como um entre um conjunto de ativos e que sua demanda se motiva como qualquer outro ativo. Como corolário principal para a política monetária, pode-se esperar que seus efeitos operem mais do que se suporia através dos efeitos diretos das mudanças no estoque de moeda sobre os gastos e menos através dos efeitos indiretos das taxas de juros no investimento e, conseqüentemente, na renda.

2.2 - O papel da política monetária

No seu discurso de posse presidencial na *American Economic Review*, Friedman (1968), atacando as políticas keynesianas das décadas de 50 e 60, seleciona duas

limitações para a política monetária que estão estreitamente relacionadas com sua concepção, já vista, de demanda por moeda e com seu enfoque sobre o comportamento da oferta agregada, que será visto adiante. Segundo o autor, a política monetária só poderá determinar as taxas de juros e a taxa de desemprego em períodos bastante limitados.

Quanto à primeira limitação, Friedman critica a suposição de que um aumento na quantidade de moeda possa afetar a taxa de juros. Diz que essa suposição parte de uma função de preferência pela liquidez negativamente inclinada, onde as pessoas seriam induzidas a guardar uma maior quantidade de moeda, ao serem diminuídas as taxas de juros. Mas adverte que a taxa mais rápida de crescimento monetário estimulará o dispêndio, a renda crescente elevará a preferência pela liquidez e a procura por empréstimos, devendo, também, elevarem-se os preços — o que reduzirá a verdadeira quantidade de moeda. Esses efeitos, segundo Friedman, inverteriam a pressão decrescente inicial sobre as taxas de juros de uma maneira bastante imediata — estipulada por ele em menos de um ano. Um outro efeito, no entanto, faria com que uma taxa mais alta de expansão monetária produzisse um nível mais alto e não mais baixo de taxas de juros. À medida que o crescimento monetário produza preços em ascensão e que o público venha a esperar que os preços continuem a subir, os emprestadores exigirão taxas de juros mais elevadas. Desta feita, Friedman argumenta que, se a autoridade monetária desejasse assegurar taxas de juros mais baixas, teria que fazê-lo por uma política monetária que, inicialmente, ao reduzir o crescimento da oferta de moeda, aumentasse a própria taxa de juros.

A segunda limitação relaciona-se à concepção de Friedman sobre a Curva de Phillips. O autor parte da suposição de que sempre haverá um nível de desemprego que possua a propriedade de ser consistente com a estrutura das taxas de salários reais — que corresponde à "taxa natural de desemprego", aquela que seria obtida por um sistema walrasiano de equações de equilíbrio geral. A análise, até então aceita, da Curva de Phillips teria a falha de não distinguir a dinâmica dos salários nominais em relação à dos salários reais.

Segundo Friedman, a autoridade monetária poderia aumentar a renda e o dispêndio através de um maior crescimento monetário. No início, ainda não esperando o aumento de preços, os trabalhadores aceitariam trabalhar pelo mesmo salário nominal, mas não mais o fariam quando observassem que os preços tivessem aumentado. A fim de manter o desemprego no seu nível mais baixo, a autoridade monetária teria de elevar ainda mais o crescimento monetário, sendo que a formação de expectativas quanto à evolução dos preços seguiria uma regra extrapolativa (adaptativa) do tipo

$$P_t^* = P_{t-1} + \alpha(P_{t-1} - P_{t-1}^*) \quad (16)$$

onde

P_t^* = nível de preços esperado no período t ;

P_{t-1} = nível de preços observado no período $t-1$;

P_{t-1}^* = nível de preços esperado no período $t-1$.

O comportamento da oferta agregada, baseado na equação (16), abrigava espaço para que a política monetária pudesse manter o desemprego abaixo de sua taxa natural desde que, ao acelerar a taxa de expansão monetária e, por conseqüência, os preços — note-se que este último processo causal é fruto da concepção friedmaniana de demanda por moeda e não uma conseqüência da estrutura da oferta agregada —, provocasse um erro provisório nas expectativas dos contratos firmados, em especial naqueles baseados em um salário real esperado. Dessa forma, o desvio em relação a uma Curva de Phillips vertical no longo prazo exigiria uma constante aceleração na taxa de expansão monetária e na taxa de inflação, o que certamente levaria a uma rota explosiva dos preços.

Como a autoridade monetária não teria controle sobre variáveis reais como a taxa de juros real e a taxa de desemprego, ao menos de forma duradoura, melhor seria, segundo a prescrição friedmaniana, que a autoridade monetária se concentrasse no controle do nível de preços. Para evitar que a moeda se transformasse numa fonte de distúrbio econômico, propunha-se a adoção de uma meta anunciada de crescimento do agregado monetário que melhor se adequasse a uma função de demanda por moeda estável. Essa taxa, que produziria uma quantidade ótima de moeda no sistema, deveria acompanhar o crescimento do produto real e a elasticidade-renda da demanda por moeda.

3 - A revolução das expectativas racionais e o monetarismo novo-clássico

O esquema adaptativo na formação de expectativas implicava que o governo poderia manter um nível de desemprego abaixo da taxa natural, desde que acelerasse a taxa de inflação em relação àquela esperada pelos agentes em função de valores observados no passado. A crítica que surge a essa suposição de comportamento era de que os agentes poderiam ser sistematicamente "enganados" na sua formação de expectativas, pois não levariam em consideração variações presentes na política governamental, como a taxa de expansão monetária, mesmo que essas informações estivessem disponíveis e fossem reconhecidas como importantes na determinação do nível de preços.

Ainda que a formulação original da hipótese de expectativas racionais seja atribuída a Muth (1961) — onde ele afirmou que, no contexto de um modelo, as expectativas dos indivíduos são racionais quando são idênticas às predições desse modelo —, foi a partir dos trabalhos de Lucas e Rappin (1969), Lucas e Prescott (1971) e Lucas (1972, 1973) que a utilização daquela hipótese em modelos macroeconômicos ganhou grande prestígio. As expectativas racionais são freqüentemente associadas ao mundo da economia novo-clássica, mas diversos trabalhos da corrente novo-keynesiana⁵ também incorporam essa

⁵ Para uma resenha destes, ver Klammer (1983) e Blanchard (1992).

hipótese em seus modelos. Na verdade, a "economia" novo-clássica inclui outras hipóteses mais restritivas, como o *clear market* walrasiano e a concorrência perfeita, abrangendo um campo amplo de pesquisas, que vai desde os primeiros modelos monetários de Lucas (1973, 1975, 1976), Sargent e Wallace (1975), Sargent (1976), Barro (1976, 1977, 1978, 1979) e Barro e Rush (1980) até as pesquisas mais recentes do Real Business Cycle e os modelos de crescimento endógeno de Paul Romer⁶.

Vale dizer, que o nosso foco de interesse nesta seção é a abordagem do conjunto inicial de trabalhos citados, que dão uma nova feição à concepção monetarista de política monetária, onde é suposta a neutralidade da moeda, inclusive no curto prazo. A limitação do espaço para que a gestão da oferta de moeda possa ter efeitos reais no nível de produto e de emprego pode ser compreendida através da curva de oferta agregada de Lucas (1973) e dos diversos trabalhos de Barro, onde se dividem os efeitos da política monetária entre os da oferta de moeda antecipada e os da não antecipada pelos agentes econômicos.

3.1 - A curva de oferta de Lucas

O entendimento dos efeitos macroeconômicos das oscilações na oferta de moeda, dentro do que se chama Monetarismo Mark II, pressupõe, por ordem lógica, a exposição de um conjunto de hipóteses e conclusões que está relacionado a uma particular compreensão da estrutura comportamental da oferta agregada. Tendo como ponto de partida o artigo de Lucas e Rapping (1969) e Lucas (1972), deve-se, ao modelo sugerido por Lucas (1973), a base para uma seqüência de trabalhos naquela linha.

Dentro de um conjunto mais amplo de modelos, onde inclui Friedman (1968), Lucas caracteriza três pressupostos particulares na sua versão:

- a) dado que o produto nominal é determinado pela demanda agregada, a divisão entre a variação do produto real e a do nível de preços dependerá do comportamento dos ofertantes de bens e trabalho;
- b) a formação de expectativas sobre os preços desconhecidos é feita de forma ótima ou racional. Ou seja, utiliza-se todo o estoque disponível de informações relevantes esperando-se um erro de previsão igual a zero. O erro observado será independente das informações disponíveis e não estará correlacionado com erros passados nas previsões;
- c) qualquer lentidão no ajustamento de curto prazo na oferta resulta da falta de informações em alguns preços relevantes para os ofertantes.

⁶ Novamente se aconselha Blanchard (1992), para uma exposição mais abrangente.

Deve-se acrescentar que, apesar de o item (c) sugerir uma forma de rigidez de preços, o modelo de Lucas supõe o *clear market* em todos os mercados.

O modelo de Lucas descrito em (a), (b) e (c) apresenta uma função de oferta agregada do tipo

$$Y_t = k_t + \alpha (P_t - P_t^*) + \lambda Y_{t-1} \quad (17)$$

onde

$$P_t^* = E_{t-1} (P_t | I_{t-1}) \quad (18)$$

sendo que

Y = produto real;

k = termo de crescimento secular, que representa progressos tecnológicos e crescimento da população;

P_t = nível de preços observado no período t ;

P_t^* = nível de preços esperado para o período t ;

I_{t-1} = conjunto de informações disponível em $t-1$.

Em comparação com o padrão de formação de expectativas descrito na equação (16), a equação (18) permite adicionar o conjunto de preços observados no período em que se formam as expectativas. Esta última equação possui uma natureza estocástica, captada por α na equação (17), já descrita no item (b). Dessa forma, a trajetória de longo prazo do produto real, determinada por variáveis reais, apenas poderá ser alterada no curto prazo quando P_t for diferente do P_t^* .

Apesar de a distribuição de probabilidade do erro de previsão do nível de preços para cada ofertante ser do tipo normal, portanto, com média igual a zero, há a possibilidade de ofertantes individuais "confundirem" os seus preços relativos observados com o nível geral de preços. Para que o produto real pudesse ser expandido no curto prazo, P_t teria de ser maior que P_t^* , ou seja, deveria haver, em determinado período, uma inflação não antecipada pelos agentes econômicos.

Ao realizar um conjunto de testes para analisar os movimentos de preços e produtos em 18 países, Lucas (1973) ressalta que seu objetivo principal era observar os efeitos de políticas de gerenciamento da demanda agregada sobre os movimentos conjuntos das taxas de inflação e produto, ou, alternativamente, de desemprego e inflação:

"The conventional Phillips curve account of this observed co-movements says that the terms of the tradeoff arise from relatively stable structural features of the economy, and thus independent of the nature of the aggregate demand policy pursued. The alternative explanation of the same observed trade off is that the positive association of prices changes and output arises because suppliers misinterpret general price movements for relative price changes" (LUCAS, 1973, p.333).

[*"A curva de Phillips convencional relacionada à esses observados co-movimentos diz que os termos do 'trade off' são explicados por características estruturais relativamente estáveis da economia, e, dessa forma, independentes da natureza da política de demanda agregada perseguida. A explicação alternativa para o mesmo 'trade off' observado é que a associação positiva das mudanças de preços e produto ocorrem porque os ofertantes confundem os movimentos no nível geral de preços com os dos preços relativos."*]

Apesar de reconhecer alguns problemas na construção estatística, como a mensuração do "produto normal" ou "pleno emprego", Lucas julga que o maior fenômeno previsto pela *natural rate theory* pôde ser comprovado pelos seus testes: quanto maior a variância da demanda agregada, mais desfavoráveis são os termos do *trade off* na Curva de Phillips, como teriam demonstrado os casos da Argentina e do Paraguai em comparação com outros 16 países.

Deve ser observado que a conhecida verticalização da Curva de Phillips é consistente com o modelo de Lucas apenas para o caso onde os ofertantes interpretam corretamente os movimentos de preços relativos. No entanto, mesmo havendo espaço para que movimentos no nível geral de preços se revertam em aumento das diversas produções individuais e do produto real total, essa possibilidade é rapidamente diminuída, à medida que se abuse dessa alternativa.

3.2 - A neutralidade da oferta de moeda antecipada

O comportamento dos agentes econômicos descrito em Lucas (1973), como da mesma forma em Sargent e Wallace (1975), dá suporte à construção de um tipo de modelo que Lucas (1975) classifica como *an equilibrium model of business cycle*. A partir de hipóteses onde os preços e quantidades são determinados, a qualquer momento, através de um processo de equilíbrio competitivo em concorrência perfeita, as expectativas são racionais e nenhum agente é perfeitamente informado sobre o estado atual da economia, geram-se resultados que mostram um tipo de economia onde "(...) real output undergoes serially correlated by movements which are not explainable by movements in the availability of factors of production" (LUCAS, 1975, p.17). [*"(...) o produto real segue serialmente correlacionado por movimentos que não são explicados por movimentos na disponibilidade dos fatores de produção."*]

Esse tipo de movimento do produto é gerado por choques monetários e fiscais não sistemáticos, que têm seus efeitos distribuídos ao longo do tempo devido a defasagens no recebimento de informações. O elo entre esse tipo de choque e movimentos reais do produto é a incorreta percepção do movimento dos preços descrito na função de oferta, na seção anterior.

No contexto acima descrito, como poderiam ser analisados os movimentos conjuntos da oferta de moeda com o produto real ou o desemprego? Robert J. Barro tenta

construir essa análise numa seqüência de trabalhos (1976, 1977, 1978, 1979, 1980), onde divide o crescimento da oferta de moeda em parcelas antecipadas e não antecipadas. A proposição central que emerge nesses textos é de que apenas a variação da oferta monetária não antecipada pelos agentes econômicos terá efeitos reais sobre o nível de produto e de emprego no curto prazo. A parcela de variação antecipada da oferta de moeda será neutra para qualquer período.

O modelo de Barro parte da definição de uma equação para a parte sistemática do crescimento monetário, do tipo

$$DM_t = \alpha_0 + \alpha_1 DM_{t-1} + \alpha_2 DM_{t-2} + \alpha_3 FED_t + \alpha_4 UN_{t-1} \quad (19)$$

onde

$$DM_t \equiv \log(M)_t - \log(M)_{t-1} \quad (20)$$

$$FED_t \equiv \log(FED)_t - [\log(FED)]_t^* \quad (21)$$

$$[\log(FED)]_t^* = \beta [\log(FED)]_t + (1-\beta) [\log(FED)]_{t-1}^* \quad (22)$$

$$UN_{t-1} \equiv \log[U/(1-U)]$$

e

M_t = média anual do estoque monetário em termos do volume de papel-moeda em poder do público mais depósitos à vista nos bancos comerciais;

FED = valor real dos gastos do governo federal;

$[\log(FED)]_t^*$ = valor "normal" dos gastos do governo federal em termos reais;

U = taxa anual média de desemprego da força total de trabalho (incluindo pessoal militar).

A parte não sistemática ou não antecipada é concebida como a diferença entre os valores observados para DM e os valores estimados DM a partir dos estimadores calculados para a equação (19), no período 1941-73, na economia americana

$$DMR \equiv DM - \widehat{DM} \quad (23)$$

onde a parcela não antecipada da oferta de moeda nada mais é do que a série de resíduos gerada pelas estimativas da oferta antecipada de moeda através da equação (19).

A partir dessas formulações iniciais, Barro testa os efeitos das parcelas antecipada e não antecipada de moeda sobre o nível de emprego e de produto real, com dados anuais (1976, 1977) e trimestrais (1978, 1979, 1980). Os testes de significância das estimativas dos parâmetros associados a DM e a DMR nas equações do nível de emprego e de produto real levaram Barro a concluir que apenas DMR influia de forma

significativa. Do ponto de vista da política econômica, resultava a confirmação de que a trajetória temporal da taxa de desemprego e do produto real não poderia ser alterada por respostas sistemáticas da oferta de moeda.

O que teria sido oferecido de novo pelo Monetarismo Mark II em acréscimo às sugestões de Friedman? Segundo Lucas (1980), não muito no que se refere à prescrição de uma regra anunciada e estável para a expansão monetária, mas a idéia de reações do setor privado contextualizadas numa formação racional de expectativas teria reforçado uma forma de pensar sobre o papel da política econômica:

"(...) as Friedman argued we should, as a problem in selecting stable, predictable policy rules. (...) our ability as economists to predict the responses of agents rests, in situations where expectations about the future matter, on our understanding of the stochastic environment agents believe themselves to be operating in" (LUCAS, 1980, p.205).

["(...) como Friedman argumentou, devemos, como um problema na seleção de regras estáveis e previsíveis de política (...) concentrar nossa habilidade como economistas para prever as respostas dos agentes, apoiando-nos, em situações onde as expectativas sobre o futuro importam, no nosso entendimento, sobre o ambiente estocástico do qual os agentes acreditam estar participando."]

4 - Considerações finais

O corpo lógico dos fundamentos da política monetária que emerge nos textos analisados parece repousar sobre dois sustentáculos: (a) supõe-se que haja uma demanda real por moeda estável, no sentido proposto por Friedman; e (b) supõe-se que os agentes privados rapidamente interpretem que a aceleração do crescimento monetário elevará a taxa de aumento do nível geral de preços e, respondendo a essa expectativa, elevem rapidamente seus preços individuais, ao invés da quantidade ofertada. Os limites impostos por (a) e (b) definem a recomendação para que a autoridade monetária não exorbite do seu papel estabilizador dos preços.

Deve-se ressaltar que a reconstrução da TQM e o comportamento da demanda por moeda analisados por Friedman — onde a taxa de juros não influencia significativamente a demanda por moeda e o crescimento do estoque monetário causa o crescimento dos preços — são implicitamente assumidos por Lucas e Barro, visto a forma como as funções de reação do setor privado interpretam o crescimento monetário. Cabe mencionar que a oposição a essa versão de demanda por moeda abrange as críticas de Tobin e Baumol nas décadas de 50 e 60, em relação à sensibilidade — juros da demanda por moeda, as análises mais recentes da corrente pós-keynesiana sobre a endogenia da oferta de moeda e o questionamento sobre a possível inversão do processo de causalidade entre a oferta de moeda e o nível de preços, inicialmente sustentado por Kaldor na década de 50.

No que se refere às conclusões extraídas por Lucas e Barro de seus modelos, há uma forte contrariedade dentro do próprio *mainstream*, atualmente disputado por novos-clássicos e novos-keynesianos. As críticas vão desde a suposição de que exista o equilíbrio geral walrasiano até as técnicas econométricas empregadas, principalmente nos trabalhos de Barro. Fora do *mainstream*, a rejeição à versão monetarista novo-clássica é ainda mais forte, envolvendo, por exemplo, a discordância dos pós-keynesianos com o tratamento probabilístico dado à "incerteza" na modelagem das expectativas.

Bibliografia

- BARRO, Robert J. (1976). Rational expectations and role of monetary policy. **Journal of Monetary Economics**, v.2, p.1-32, May.
- BARRO, Robert J. (1977). Unanticipated money growth and unemployment in the United States. **American Economic Review**, v.67, p.101-115, Mar.
- BARRO, Robert J. (1978). Unanticipated money out put and the price level in the United States. **Journal of Political Economy**, v.86, p.549-580, Aug.
- BARRO, Robert J. (1979). More on unanticipated money growth. **American Economic Review**, v.69, p.327-343, Jul.
- BARRO, Robert J., RUSH, M. (1980). Unanticipated money and economic activity. In: FISCHER, Stanley, ed. **Rational expectations and economic policy**. Chicago: The University of Chicago.
- BLANCHARD, Olivier J., FISCHER, Stanley (1989). **Lectures on macroeconomics**. Londres: The MIT.
- BLANCHARD, Olivier J. (1992). Novos clássicos e novos keynesianos: a longa pausa. **Literatura Econômica**, Rio de Janeiro: IPEA, n. esp., p.16-30, jun.
- FISHER, Irving (1911). **The purchasing power of money**. New York: Macmillan.
- FRIEDMAN, Milton (1956). The quantity theory of money: a restatement. In: _____, ed. **Studies in the quantity theory of money**. Chicago: The University of Chicago.
- FRIEDMAN, Milton (1959). The demand for money: some theoretical and empirical results. **Journal of Political Economy**, v.67, p.327-351, Aug.
- FRIEDMAN, Milton (1968). The role of monetary policy. **American Economic Review**, v.58, p.1-17, Mars.
- FRIEDMAN, Milton (1970). A theoretical framework for monetary analysis. **Journal of Political Economy**, v.78, p.195-238, Mar./Apr.

- FRIEDMAN, Milton, SCHWARTZ, A. J. (1963). **A monetary history of the United States, 1867-1960**: National Bureau of Economic Research. Princeton: Princeton University..
- FRISCH, Helmut (1983). **Theories of inflation**. Cambridge: Cambridge University.
- KEYNES, John M. (1936). **The general theory of employment, interest, and money**. London: Macmillan.
- KLAMER, A. (1983). **Conversas com economistas**. São Paulo: Pioneira.
- LAIDLER, David (1978). **The demand for money**. New York: Dun-Donneley.
- LUCAS, Robert E. (1972). Expectations and the neutrality of money. **Journal of Economic Theory**, v.4, p.103-124, Apr.
- LUCAS, Robert E. (1973). Some international evidence on output inflation trade-offs. **American Economic Review**, v.63, p.326-334.
- LUCAS, Robert E. (1975). An equilibrium model of the business cycle. **Journal of Political Economy**, v.83, p.1113-1144, Dec.
- LUCAS, Robert E. (1980). Rules, discretion, and the role of the economic advisor. In: FISCHER, Stanley, ed. **Rational expectations and economic advisor**. Chicago: The University of Chicago.
- LUCAS, Robert E. (1976). Econometric policy evaluation: a critique. In: BRUNNER, Karl, MELTZER, Allan, ed. **The Phillips curve and labor markets**. Amsterdam: North Holland.
- LUCAS, Robert E., RAPPING, Leonard A. (1969). Real wages employment and inflation. **Journal of Political Economy**, v.77, p.721-754, set.
- LUCAS, Robert E., PRESCOTT, Edward C. (1971). Investment under uncertainty. **Econometrica**, v.39, p.721-754, set.
- MORAES JÚNIOR, Aod C. (1993). Controle da oferta de moeda: monetaristas, keynesianos neoclássicos e pós-keynesianos. In: ENCONTRO NACIONAL DE ECONOMIA, 21., Rio de Janeiro. **Anais...** Rio de Janeiro: ANPEC.
- MORAES, Roberto C. (1991). As origens da teoria quantitativa da moeda e de seus primeiros críticos. **Perspectiva Econômica**, v.69, p.75-94, abr./jun.
- MUTH, John F. (1961). Rational expectations and the theory of price movements. **Econometria**, v.29, p.133-169, maio.
- SARGENT, Thomas, (1976). Rational expectations and the theory of economic policy. **Journal of Monetary Economics**, v.2, p.169-183, Apr.
- SARGENT, Thomas, WALLACE, Neil, (1975). Rational expectations, the optimal monetary instrument, and the optimal money supply rule. **Journal of Political Economy**, v.83, p.241-255, Apr.

SELDENS, Richard T. (1956). Monetary velocity in the United States. In: FRIEDMAN, Milton. **Studies in the quantity theory of money**. Chicago: The University of Chicago.

TRICHES, Divanildo (1990). **Demanda por moeda no Brasil e a causalidade entre as variáveis monetárias e a taxa de inflação: 1972-1987**. Porto Alegre: UFRGS/CPGE. (mimeo)

Abstract

The purpose of this article is to show the evolution of the theoretical structure which gives support to the monetarist monetary policy guide. The article analyses the principal contributions developed by Milton Friedman, Robert Lucas Jr. and Robert Barro on the different stages of the monetarist view in order to point out particular hypothesis, about the demand for money and the aggregate supply behavior.