Search Frictions in Financial Markets and Monetary Policy *

Scott Hendry
Bank of Canada

Kevin Moran[†]
Bank of Canada

March 2004. *Preliminary and Incomplete*

Abstract

This paper introduces search frictions in financial markets, within a standard quantitative monetary model, with the objective of generating greater persistence in the effects of monetary policy shocks. Specifically, the paper first assumes that banks enter into contact with suitable new entrepreneurs-clients via a search-and-matching mechanism similar to that used to study labour markets (Pissarides, 1985). Here, the interest rate on loans implements the sharing of the surplus associated with each bank-entrepreneur match. Second, we assume that the entrepreneurs, who oversee the production of the economy's output, operate a diminishing-returns-to-scale technology.

In this environment, banks have an incentive to use part of any unexpected liquidity injection to search for new clients rather than lending it all out to its existing ones, because spreading out a given supply of funds across a wide pool of projects attenuates the effects of diminishing returns to scale. Further, these increased search efforts generate persistent changes in the relative bargaining power of banks and entrepreneurs, which affects surplus sharing and thus the lending rate that implements it. The combination of these two effects is shown to add persistence to the decrease in nominal interest rates that follows monetary policy easings and to create hump-shaped responses in output and inflation.

JEL classification: E4, E5

*We thank Christian Calmès, Merwan Engineer, Allen Head, Peter Ireland, Césaire Meh and Shouyong Shi for useful discussions and suggestions. We also benefitted from the comments of participants at the 2002 annual meeting of the Canadian Economics Association as well as the 2003 meeting of the European Economics Association, and seminar participants at Laval University, Simon Fraser University, and the University of Victoria. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

[†]Corresponding author. Email: kmoran@bankofcanada.ca. Address: MFA Department, Bank of Canada, 234 Wellington, Ottawa On. K1A 0G9; Canada.

1. Introduction

This paper incorporates search frictions in financial markets, within the limited participation class of quantitative, monetary models, and verifies whether this extension generates greater persistence in the responses of economic variables following monetary policy shocks.

The two main assumptions guiding the construction of the model are the following. First, the entrepreneurs, who oversee the production of the economy's output and transact with banks to secure financing of their activities, operate a diminishing-returns-to-scale technology. Second, finding new clients to which to lend is neither instantaneous nor costless for banks but instead arises from search activities.

Operating in such a world, banks are reluctant to make available to their existing clients the full amount of loanable funds' injections by monetary authorities. Doing so would push these clients into less profitable areas of their technology and hence reduce returns from lending. Instead, banks choose to direct some of these additional funds towards the search for new clients. In time, this search expands the pool of entrepreneurs linked to the bank, allowing it to spread available funds more efficiently and escape the diminishing returns to scale faced, individually, by its clients.

The process by which banks enter into contact with new clients resembles closely those featured in the search-and-matching labour literature.[1] There exists a pool of entrepreneurs, each endowed with a project that they must finance externally. Some of these entrepreneurs have an ongoing lending relationship with a bank and can thus negotiate a lending contract and produce. Others do not and search in order to establish such a relationship. Banks open 'lines of credit' backed by liquid assets which are to be lent out if they become matched with an entrepreneur. The matching technology takes as inputs the size of the pool of searching entrepreneurs and the real value of liquid assets held by banks; its output consists of new matches between banks and entrepreneurs.[2] Once a bank-entrepreneur pair has been formed, the two participants jointly decide the scale

[1]Following the early contribution of Pissarides (1985), a large literature has developed to study the implications of the search-and-matching environment for labour market issues. Shimer (2003) presents an overview and critical assessment of this literature. This environment has been embedded into quantitative models of the RBC lineage by Merz (1995), Andolfatto (1996), and Cooley and Quadrini (1999a).

[2]The pool of entrepreneurs thus corresponds to the unemployed workers of the labour literature, whereas the lines of credit opened by banks correspond to vacancies advertised by firms.

of production that maximizes the net surplus accruing from the match and then share that surplus. The nominal lending rate on the loan extended to the entrepreneur is the instrument by which this sharing is implemented. Consequently, this rate may react to factors like the relative bargaining power of the two match participants.

Compared to the standard limited participation framework, a monetary easing has, in our model, the following set of effects. First, banks do not lend out the entire liquidity injection but rather set aside a fraction of it to search and attract new entrepreneurs-borrowers. Less funds are thus initially circulated into the economy, whith the potential to induce similar delays in the response of inflation to the shock. Second, the increased search efforts by banks increase the probability that an unmatched entrepreneur succeeds in establishing a lending relationship. This improves the outside option (and the bargaining power) of entrepreneurs presently matched with a bank; as a result, banks need to reduce their lending rates. In subsequent periods, the increased search efforts by banks have increased the number of matched entrepreneurs. The pool of those still searching is thus reduced, which makes it easier for them to find a match. This keeps the outside option of matched entrepreneurs high and continues to apply downward pressure on the lending rates and intermediation spreads banks can demand. In the findings presented here, these two channels appear operational: they add persistence to the decrease in nominal interest rates and they generate hump-shaped responses in output and inflation. Further analysis will explore the robustness of those findings.

The rest of this paper is arranged as follows. Section 2 discusses the contribution of this paper, first to the limited participation class of monetary models and, second, to the more general literature on credit market imperfections. Section 3 presents the model, while section 4 discusses the calibration and solution strategies used to provide quantitative content to its implications. Section 5 presents those results: it explores the impact of monetary policy and technology shocks, as well as shocks to the matching process. Section 7 concludes.

2. Context and Related Literature

2.1 Limited participation models

In the standard limited participation model (e.g., Christiano (1991)), a decrease in nominal interest rates (the liquidity effect) results from a monetary easing because the banking sector, faced with an unexpected injection of loanable funds, lowers lending rates in to entice firms to borrow all the additional available funds. Firms take advantage of these unexpected low costs of borrowing to expand their operations, increasing their labour demand and overall production.

The liquidity effect's origin lies in the relative abundance of funds in the banking sector relative to those in the goods sector. Households are precented, by assumption, from correcting this imbalance during the impact period (which they could do by transferring some of their deposits back to the goods sector). The crucial role of this imbalance explains why it is difficult to generate persistence in the liquidity effect. As long as newly injected liquidity makes its way back to households (in the form of wage or dividend payments), households hold all the economy's cash stock at the end of the impact period. They are thus able to correct the initial imbalance and send optimal amounts of liquidity to the goods and financial sectors by the beginning of the next period.

The search for persistence has most often been centered around the decision on the households' side. For example, making it costly to adjust portfolios can prevent households from correcting the imbalance. In such an environment, households continue to send relatively large amounts of liquidity to banks in order avoid the adjustment costs and thus the relative abundance of loanable funds in the banking sector persists.[3] Alternatively, Andolfatto and Gomme (2003) and Andolfatto et al. (2002a,b) assume that households have incomplete information about the nature and the duration of the monetary shock. They continue to send 'too much' liquidity to the banking sector because of incorrect expectations; persistence now depends on the speed of learning.

[3] Recent papers using limited participation with portfolio adjustments costs include Cooley and Quadrini (1999a) and Christiano and Gust (1999). Models that refine the timing of events in a given period, like that of Alvarez and Atkeson (1997), could also be interpreted as models with portfolio adjustment costs that are measured in time.

Turning instead to the banking sector to find a mechanism to increase persistence, as the present paper does, echoes the strategy employed by Cook (1999), in which it is assumed that the banks' intermediation costs are lowered by past lending activity.[4] The initial lending increase implied by the liquidity injection thus makes lending less costly in future periods and the lending rate (but not the deposit rate) returns only gradually to its pre-shock level. More generally, the present paper shares with Chari et al. (1995) the belief that the limited participation framework requires a more sophisticated banking sector to achieve its potential as a complete representation of the monetary transmission mechanism.[5]

Assuming that there are diminishing returns to scale in the production of goods would not, by itself, create a persistent liquidity effect. It simply would make firms more reluctant to expand their operations and borrow the additional liquidity, forcing banks to lower even further the lending rate until a new equilibrium is attained. The joint hypothesis of diminishing returns and costly search, on the other hand, offers bank an alternative use of their funds, because holding liquid assets rather than lending them out helps increase the flow arrival of new clients. In turn, a greater number of clients allows the bank to spread available funds more efficiently and thus increase the returns from lending.

2.2 Search and credit markets

Perhaps the main function of financial intermediaries is to alleviate, if not overcome, information frictions or asymmetry that exist between borrowers and lenders. The literature on the ways in which intermediaries accomplish such a task is very large. One notable subset of that literature is the development, using models from the 'New-Keynesian' class, of the aggregate implications for monetary policy of such information frictions: see the contributions of Bernanke et al. (1999) and Carlstrom and Fuerst (2001).

Another type of friction that financial intermediaries might also help to overcome is what Becsi et al. (2000) refer to as "search and entry" frictions. These may represent

[4] Specifically, the technology used by banks to verify the credit worthiness of borrowers. The assumption is similar to the learning-by-doing mechanism often used in growth models.

[5] The model in Chari et al. (1995) possesses a sophisticated banking sector that includes a technology for deposit creation and the distinction between currency, the monetary base and broader monetary aggregates. However, it relies on portfolio adjustment costs to induce persistence in the liquidity effect.

costs related to the initial participation of firms into loan markets, the costs involved in negotiating lending contracts, etc. It is this type of friction that the present paper seeks to explore.

We envision a world where entrepreneurs are endowed with a project, but no recognizable collateral. Entrepreneurs that already have a lending relationship with a bank can get financing and produce. Entrepreneurs that are not presently engaged in such a relationship search for one and meet with representatives from banks (hereafter, we call them loan officers), one at the time, and explain to them the nature of their project or idea. Whether a loan is extended then depends on the intangible features of the meeting between the entrepreneur and the loan officer. One loan officer will find the project proposed by an entrepreneur feasible, while another will judge the same idea as too risky. It is in that way that we interpret in our environment the random nature of the matching function and the undirected search it implies.

The lending market we model thus reserves a key role to the relationship between the loan officer and the entrepreneur, in a world where information is imperfect and collateral largely absent. This implies that our paper is probably a better vehicle for the analysis of lending to small and medium businesses than lending to large corporations. Accordingly, the quantitative results presented in section 5, detailing the effects of monetary policy shocks on our artificial economy, should be understood as counterparts to the effects on small businesses from actual monetary policy shocks.[6]

Some recent papers also explore the consequences of using the search and matching framework to analyze credit markets (Wasmer and Weil (2000), Becsi et al. (2000, 2002), Den Haan, Ramey and Watson (1999)). However, the model presented here is the first to include those frictions in a fully quantitative and dynamic framework in which the aggregate implications of these frictions and the manner that they impinge on monetary policy can be studied.

To keep the analysis simple, we assume that the loans extended by the banks to the entrepreneurs, once matches are formed, are risk-free as no uncertainty is revealed during

[6]Gertler and Gilchrist (1994) report that the effects of monetary policy shocks depend on the size of firm. Cooley and Quadrini (1999b) present a quantitative, structural model that is consistent with the facts identified by Gertler and Gilchrist.

the lifetime of the loan. Although we abstract from information asymmetry, one can easily imagine a model where first the search and entry frictions we emphasize, and then the information frictions modelled in Bernanke et al. (1999) and Carlstrom and Fuerst (2001) would be introduced; the resulting model would present a rich description of the role financial intermediaries play in the transmission of monetary policy shocks.

3. A Simple Model

In this section, we present a simple model in which the search for new clients that banks conduct is a one-period activity. The liquid assets that a bank chooses to hold (rather than lend out) during the period increase the chance of finding new clients; the opportunity cost of that activity is thus the lending rate banks charge to their existing clients. There are no other costs to search and at the end of the period, the liquid assets that were allocated to search activities are still available and can be returned to households.[7] This simple framework helps showcase clearly the environment we envision. We plan to later develop a model where search activities can require banks to hold liquid assets for longer periods of time.

3.1 The central bank

At the beginning of each period, the central bank deposits an amount of liquidity (loanable funds)[8] in the banking sector, on behalf of the households. The (net) growth rate of these monetary injections follows a simple exogenous and stochastic process such as the following AR(1):[9]

$$x_{t+1} = (1-\rho)x + \rho x_t + \varepsilon_t^m, \quad \varepsilon_t^m \sim (0, \sigma_m^2).$$

[7]No money is thus lent out in the period where search takes place. The requirement that liquidity must be kept idle in order to help the searching efforts of banks can be interpreted as representing the differing frequencies of the process by which banks make contact with potential clients and the one by which they extend loans. If a given monetary injection is completely lent out for a relatively long period of time, any match that is formed during the period of that loan will be idle because of the absence of liquidity in the bank's vaults.

[8]Discuss the equivalence between currency and loanable funds: assumption of no inside money creation, or at least of a stable multiplier.

[9]Note that recent work by Christiano and Eichenbaum (1998) has established that such an AR(1) process is a good approximation to the available empirical evidence only when M1 is the monetary aggregate one wishes to match to the model's concept of money. A low order MA process better fits the available evidence when M2 is considered the real world counterpart to money.

Note that while this money is deposited in their behalf, households cannot transfer it back to the goods market in the current period, even if they wish to do so. This is the limited participation constraint.

3.2 The banking sector

At the start of a given period, a typical large bank has in its disposition an amount of liquidity $M_t^d + X_t$, deposited either directly by households or on their behalf by the central bank. This bank can also count on working relationships with a fixed number of entrepreneurs. The bank must decide on the allocation of these funds between the borrowing needs of its existing clients and the search for new ones. This allocation thus obeys the following constraint:

$$M_t^d + X_t = B_t + S_t,$$

where B_t represents the total lending by this bank to existing clients and S_t the amount directed towards the search for new clients. Those latter amounts are left unused in the current period: instead of flowing through the economy in the form of credit to its existing clients, they are dispatched to loan officers (see below) to be ready to be lent if the officers are successful in finding promising new entrepreneurs to lend to. Evidently, the higher the return a bank can generate by lending to its existing clients, the less incentive it has to keep these liquidities idle in the hope of attracting new clients.[10] At the end of the period, existing clients reimburse their loans and the bank uses these proceeds along with the idled liquidities to pay back its obligation to depositors and shareholders.

Having described the working of a large bank, we now distinguish between the behavior of two different agents in such a bank: the loan officer and bank headquarters.

3.3 Loan officers

Loan officers conduct the bargaining with the clients of the bank. A loan officer is either active, when he is matched with a suitable entrepreneur, or inactive, when even

[10] Although we assume that these liquid assets earn a zero rate of return, they would be modelled more realistically as assets with a lower return than lending. Consequently, the discussion that follows should be interpreted in terms of the excess return earned by loans.

though an amount of liquidity as been earmarked for him, he has not found any suitable entrepreneur to lend it to.[11] Note that the client of one loan officer is assumed to be unable to transfer to another officer because the quality of the match would deteriorate significantly. The metaphor of the loan officer is meant to represent elements of banking relationships (personal knowledge, confidence, etc.) that are not transferable. Active loan officers conduct business with their clients whereas inactive ones search for good matches with entrepreneurs.[12]

Denote the value function of an active loan officer by V_t^{lo}, and the value function of an inactive officer (unused line of credit) by V_t^{lc}. V_t^{lo} is determined by the following Bellman equation:

$$V_t^{lo} = \frac{\lambda_t(R_t^l - R_t^d)B_t}{P_t} + \beta\{smE_t[V_{t+1}^{lc}] + (1-sm)E_t[V_{t+1}^{lo}]\}. \tag{1}$$

.

The cash flow accruing to an active loan officer consists in its gross loan profits, $(R_t^l - R_t^d)B_t$, where R_t^l is the lending rate, B_t is the total nominal loans extended to the entrepreneur and his costs of funds is R_t^d, an internal, or shadow rate paid to bank headquarters (see below). The aggregate price level P_t is used to express this cash flow in real terms. Current cash flows are valued at λ_t, the marginal utility of income of banks shareholders. Additionally, a match with an entrepreneur carries a forward looking component: with probability sm, the match will be terminated at the end of the period and this loan officer will become inactive. With probability $1 - sm$, he will stay active.[13]

Assume that a nominal amount $P_t sc$ in liquid assets must be kept idle in order to create one new line of credit. Immobilizing money for one period carries the opportunity cost of R_t^d, paid to bank headquarters. In value terms, opening one line of credit therefore

[11]Such an officer could also be interpreted as an unused, open line of credit. In what follows we use the expressions 'inactive loan officer' and 'unused line of credit' interchangeably.

[12]Consider the following metaphor for the search process modelled here. Every Sunday, the chamber of commerce organizes a brunch where potential entrepreneurs are invited to come and meet representatives from the banks (the inactive loan officers). The entrance is free for potential entrepreneurs but loan officers must show, before given admittance, that they have the means and the authority to extend loans if they find a suitable match. In order to do so, they must prove that funds are earmarked to them (the S_t funds). If, while mingling through the room, a match is created, the entrepreneur and loan officer agree to meet again on Monday morning to discuss the terms of the loan and the entrepreneur starts production later that day.

[13]We assume that sm is exogenously determined. The possibility of making this rate endogenous is explored in section 6.

entails a trade-off between the negative current cash flow and the opening of a new credit relationship tomorrow. Accordingly, the value function for an opened line of credit (an inactive loan officer) obeys the following Bellman equation:

$$V_t^{lc} = -\lambda_t(R_t^d - 1)(P_t sc) + \beta\{\theta_t^b E_t[V_{t+1}^{lo}] + (1-\theta_t^b)E_t[V_{t+1}^{lc}]\}, \tag{2}$$

where θ_t^b represents the probability that the inactive loan officer will connect with a suitable borrower this period and become active in the next period.

We can now define the surplus enjoyed by a loan officer who is matched with a suitable entrepreneur:

$$S_t^b = V_t^{lo} - V_t^{lc}. \tag{3}$$

3.4 Bank headquarters

We assume that a typical bank consists of a large organization with many loan officers. Some of these are active while the remainder are inactive. Bank headquarters must insure that enough funds are available for all lending and searching activities of its officers and that the allocation of these funds between the two activities is efficient. It thus determines the internal rate R_t^d that will implement such an efficient allocation. In the simple model presented in this section, banks solve a one period problem; the additional assumption of free entry into the banking sector insures that a simple no-profit condition can be used to determine the internal rate R_t^d.

Insuring that enough funds are available for the activities of all loan officers entails the following budget constraint for bank headquarters:

$$M_t^d + X_t = v_t^e B_t + S_t, \tag{4}$$

where the left hand term represents total available funds and the right hand term adds up total lending (B_t to each active loan officer, of which there are a number v_t^e for each bank) and total liquidities held for search.

End of period assets of bank headquarters consist of two parts: the repayments (principal and interest) from active loan officers and the return (of 1) on liquid assets held for search purposes. End of period liabilities are simply payments on households' deposits at the (external) rate r_t^d. Profits for bank headquarters are thus:

$$\Pi_t^b = v_t^e B_t R_t^d + S_t - (M_t^d + X_t) r_t^d. \tag{5}$$

Assuming free entry so that the no profit condition holds, equations (4) and (5) lead to a relationship between the deposit rate and the internal rate: headquarters will charge their loan officers a mark-up over what it pays to its depositors to insure that it can pay for the unprofitable (from a current cash flow perspective) search activities.[14]

3.5 Production sector

There is a continuum of identical entrepreneurs in the economy, located on the interval $[0, e]$; the remaining agents in the economy, numbering $1 - e$, are workers. Each entrepreneur has access to a project that relates the labour of workers she employs and possibly her own labour (in a management responsibility) to consumption goods production. The production function has the following form:

$$F(n_t^e, n_t^w) = A_t (n_t^e)^{\alpha_1} (n_t^w)^{\alpha_2}, \quad \alpha_1 + \alpha_2 \leq 1; \tag{6}$$

where n_t^e is the entrepreneur's managerial work and n_t^w the amount of labour she hires. The identification of n_t^e with an entrepreneur's management of her own project implicitly assumes there are subtle features about each technology so entrepreneur skills are not transferable. Consequently, even if $\alpha_1 + \alpha_2 = 1$, an entrepreneur cannot expand her operations without bound because of the increasing marginal utility of her time.

The assumption of diminishing returns to scale that is present in the formulation of technology above can be interpreted as the difficulties that small business owners face when they try to replicate a successful business and try to expand. A very hands-on entrepreneur, for example, will find it difficult to monitor two or three locations as well as she monitored the first location. If this close monitoring was a key ingredient in the success of the first location, the expansion will imply reduced profitability. It is this difficulty in managing growth that the production function we imply is meant to capture.

[14]Note that the role played by the 'bank headquarter' is simply to diversify risk (the risk that any particular inactive loan officer will not find a suitable entrepreneur) and is therefore somewhat of a veil. Similar results could obtain where the households to allocate funds between active and inactive loan officers directly, as is the case in Cooley and Quadrini (1999a).

We assume that entrepreneurs do not have the funds to hire their workers and must rely on bank loans to do so. At any moment of time, only a fraction of entrepreneur are matched with a loan officer and are thus able to operate their technology. Denote this fraction by v_t^e.[15] The remaining $1 - v_t^e$ fraction of entrepreneurs are inactive and have no revenues.

Denote by J_t^e the value function of an active entrepreneur (with access to financing) and by J_t^{ne} the value function of an entrepreneur with no financing. The value function of the active entrepreneur is comprised of a current cash flow, a current utility flow and a continuation value; it is expressed by the following Bellman equation:

$$J_t^e = \frac{\lambda_t (P_t F(n_t^e, n_t^w) - R_t^l W_t n_t^w)}{P_t} - l(n_t^e) + \beta\{smE_t[J_{t+1}^{ne}] + (1 - sm)E_t[J_{t+1}^e]\}. \qquad (7)$$

At the beginning of each period, an active entrepreneur negotiates a borrowing rate, R_t^l, and the amount to borrow with her bank manager. She then hires the unskilled labour necessary to the operation of her technology. Current cash flows resulting from production are thus $P_t F(n_t^e, n_t^w) - R_t^l W_t n_t^w$, which are discounted by the marginal utility of income of the entrepreneur. Again, the aggregate price index P_t is used to express the cash flow in real terms. The current (dis)utility flow, $l(n_t^e)$, results from the labour the entrepreneur puts in managing her project. Finally, the continuation value depends on the probability that the match will be terminated at the end of the current period, in which case the entrepreneur becomes inactive.

An inactive entrepreneur has no revenues and makes no decisions in the current period. However, she searches (at no cost) for a suitable loan officer, in the hope of becoming active next period. Her value function is thus expressed as follows:

$$J_t^{ne} = \beta\{\theta_t^e E_t[J_{t+1}^e] + (1 - \theta_t^e)E_t[J_{t+1}^{ne}]\}, \qquad (8)$$

where θ_t^e expressed the probability that an inactive entrepreneur finds a loan officer during the current period and can become active in the next. Note again that there are no decision and no costs related to search and thus this probability is exogenous from the point of view of the entrepreneur. In equilibrium, this probability will be linked to the number

[15]Therefore, the economy-wide number of entrepreneurs matched with a bank at time t is $e\, v_t^e$

of inactive entrepreneurs and to the real value of liquid assets banks direct towards the search for new clients.

We can now define the surplus associated with an entrepreneur having a match with a loan officer:

$$S_t^e = J_t^e - J_t^{ne}. \tag{9}$$

3.6 Bargaining

Within a match, the scale of operations of the entrepreneur and the lending rate that is applied to her loan are decided according to the following bargaining scheme. First, define the total surplus of the match by the sum of the two individual surplus:

$$TS_t = S_t^b + S_t^e. \tag{10}$$

Introducing (3) (with (1) and (2)) and (9) (with (7) and (8)) in (10), leads to the following expression for the total surplus:

$$TS_t = \lambda_t(F(n_t^e, n_t^w) - R_t^d w_t n_t^w) - l(n_t^e) + \frac{\lambda_t(R_t^d - 1)sc(1 - sm - (1-\xi)\theta_t^e))}{\theta_t^b \xi} \tag{11}$$

Next, find the optimal scale of operation for the entrepreneur. Because it is related to the effort the entrepreneur puts in managing the project and the amount of labour she hires, this optimal scale arises as the result from the following maximization:

$$\max_{[n_t^e, n_t^w]} TS_t. \tag{12}$$

Considering the expression of the surplus in (11), this maximization leads to the following first order conditions for the choice of n_t^e and n_t^w:

$$\lambda_t F_1(n_t^e, n_t^w) = l'(n_t^e); \tag{13}$$

$$F_2(n_t^e, n_t^w) = R_t^d w_t. \tag{14}$$

Note that this last condition, expressing labour demand, is the same as the one which obtains usually in limited participation models: in that respect, the search and marching framework used here replicates the standard model.

Finally, the entrepreneur and her loan officer divide the maximized surplus according to predetermined (an exogenously given) shares, with a fraction ξ going to the bank and the remainder, $1-\xi$, to the entrepreneur.[16] The lending rate is determined as a residual of this operation: it is the rate that implements such a sharing. Considering the definition of all three surplus, the lending rate is such that the following obtains:

$$S^b = \xi\, TS_t. \tag{15}$$

Expanding on this equation (by using (11) and (3)), one is lead to the following expression solving (implicitly) for the lending rate:

$$\begin{aligned} -\xi F(n_t^e, n_t^w) + (R_t^l - (1-\xi)R_t^d)w_t n_t^w - l(n_t^e)/\lambda_t \\ + \frac{(R_t^d - 1)sc(1-\xi)\theta_t^e}{\theta_t^b} = 0. \end{aligned} \tag{16}$$

3.7 Search and matches

The flow arrival of loan officers-entrepreneurs matches depends on the number of inactive entrepreneurs, $1-v_t^e$, and the real value of liquidities directed by banks towards the search for new clients. This search is an activity defined in finite terms, so that banks open the equivalent of lines of credit. Opening one additional line of credit requires that a nominal amount of $P_t sc$ be immobilized, with sc the real unit cost. A bank thus needs $v_t^b P_t sc$ in liquid assets to open v_t^b such lines of credit. The two-side search results in m_t new entrepreneur-officer relationships, with the following functional form used:

$$m_t = m(1-v_t^e, v_t^b) = \chi_t(1-v_t^e)^\eta (v_t^b)^{1-\eta}, \tag{17}$$

where $1-\eta$ is the elasticity of match arrival with respect to search efforts by banks χ and is a measure of the productivity of the matching function.[17]

[16] This sharing rule can be formally justified by the outcome of a Nash bargaining process that maximizes the following product:

$$\max (S_t^b)^\xi (S_t^e)^{1-\xi},$$

where ξ represents the bargaining power of the loan officer.

[17] Note that this measure of the productivity of the matching process could be made stochastic in order to study the consequences of shifts in its value. Moreover, long-term trends in credit could be envisioned to arise from trends in the value of that parameter.

Recall that a proportion sm of the existing matches are lost at the end of the current period. Considering this creation and destruction of matches, the following law of motion for the fraction of entrepreneurs matched with a loan officer arises:[18]

$$v_{t+1}^{e} = (1 - sm)v_t^e + m(1 - v_t^e, v_t^b). \tag{18}$$

In equilibrium, the probabilities that a loan officer searching for a new client and an entrepreneur searching for financing find a match, respectively denoted by θ_t^b and θ_t^e, are equal to the following:

$$\theta_t^b = \frac{m(1 - v_t^e, v_t^b)}{v_t^b}; \tag{19}$$

$$\theta_t^e = \frac{m(1 - v_t^e, v_t^b)}{1 - v_t^b}. \tag{20}$$

Define the tightness of the credit market by:

$$\phi_t = \frac{1 - v_t^e}{v_t^b}. \tag{21}$$

An increase in ϕ_t, the relative number of entrepreneur searching for credit over the number of opened lines of credit, indicates a tighter access to credit for inactive entrepreneurs. Considering the probabilities in (19) and the functional form in (17), the probabilities can be rewritten as follows:

$$\theta_t^b = \chi_t \phi_t^{\eta}; \tag{22}$$

$$\theta_t^e = \chi_t \phi_t^{1-\eta}. \tag{23}$$

3.8 Households

We assume that households are risk neutral. This assumption is present because the surplus sharing rule we use is the solution of a Nash bargaining process only under risk neutrality.[19] This assumption insures that the marginal utility of income is equal across

[18] The timing of events considered here precludes the possibility that new matches get terminated before they have started being active.

[19] Cooley and Quadrini (1999a) show that the introduction of risk aversion does not modify sharply the results of their analysis. One would thus imagine that the introduction or risk aversion in our econoy would also not modify drastically our results.

workers and both types of entrepreneurs.[20] Households make choices on consumption and the allocation of their assets; workers must also choose their work effort.[21] A typical worker's decisions are subject to a cash-in-advance constraint and a budget constraint. His maximization problem is thus the following:

$$\max_{\{c_{t+k}, N^W_{t+k}, M^c_{t+k+1}, M^d_{t+k+1}\}_{k=0}^{\infty}} E_t \sum_{k=0}^{\infty} \beta^{t+k}[c_{t+k} - l(N^W_{t+k})] \tag{24}$$

with respect to:

$$c_t \leq \frac{M^c_t}{P_t}, \quad \text{t=1,2,...} \tag{25}$$

and

$$\frac{M^c_{t+1} + M^d_{t+1}}{P_t} \leq \frac{r^d_t M^d_t}{P_t} + \frac{W^d_t N^w_t}{P_t} + \frac{\Pi^b_t}{P_t}, \quad \text{t=1,2,...} \tag{26}$$

In (26), dividends from banks are denoted by Π^b_t and assumed to be paid to workers. Let μ_t be the (real value of the) Lagrangian multiplier associated with the cash-in-advance constraint (25) and λ_t be the equivalent multiplier for the budget constraint (26). The first order conditions of the problem with respect to c_t, N^w_t, M^c_{t+1}, and M^d_{t+1} are the following:

$$1 = \mu_t \tag{27}$$

$$l^{'}(N^W_t) = \frac{\lambda_t W_t}{P_t} \tag{28}$$

$$\frac{\lambda_t}{P_t} = \beta E_t[\frac{\mu_{t+1}}{P_{t+1}}] \tag{29}$$

$$\frac{\lambda_t}{P_t} = \beta E_t[\frac{\lambda_{t+1} r^d_{t+1}}{P_{t+1}}] \tag{30}$$

3.9 Definition of the competitive equilibrium

The competitive equilibrium is a solution for the following set of 19 unknowns, $\{c_t$, n^w_t, N^w_t, μ_t, λ_t, V^b_t, S^e_t, S^b_t, B_t, v^b_t, v^e_t, P_t, W_t, R^l_t, R^d_t, S_t, M^d_t, M^c_t and $r^d_t\}$. This equilibrium is such that the total surplus TS_t is maximized, the sharing of this surplus obeys (12), bank headquarters makes zero profits, households solve their maximization problems, and all the following market clearing conditions are satisfied.

[20] Alternatively, we could follow Bernanke et al. (1999) and assume that entrepreneurs are risk neutral but workers are risk averse; we would then have to keep track of two different sets of marginal utility.

[21] The decision of entrepreneurs regarding their own work effort takes place during the bargaining process with their loan officer

First, in the goods market, aggregate consumption must equal production. The latter is the product of each active entrepreneur's production and the number of such entrepreneurs:

$$c_t = v_t^e F(n_t^e, n_t^w). \tag{31}$$

In the money market, total available liquidities must be sufficient to cover aggregate lending and liquidities directed towards search:

$$M_t^d + X_t = B_t + S_t. \tag{32}$$

Within each match, the loan supplied must be sufficient to cover planned hiring:

$$B_t = W_t n_t^w. \tag{33}$$

The amount of liquidity available for search must be sufficient to cover all planned search activities by loan officers:

$$S_t = v_t^b P_t sc. \tag{34}$$

Finally, total labour demand (the amount hired by each active entrepreneur times the number of such entrepreneur) must equal total labour supply:

$$(ev_t^e)n_t^w = (1-e)N_t^w. \tag{35}$$

4. Calibration and solution

Recall that the production function was defined in a general way as:

$$F(n_t^e, n_t^w) = A_t (n_t^e)^{\alpha_1} (n_t^w)^{\alpha_2}$$

In the simplified version of the model, we assume that an entrepreneur's work does not enter the production function ($\alpha_1 = 0$) and that $\alpha_2 = 0.7$; we now have:

$$F_t = A_t (n_t^w)^{0.7}. \tag{36}$$

We assume that the elasticity of match-arrival with respect to the number of active entrepreneurs, η, is equal to 0.2. Although no empirical estimates of that parameter exist, it could be interpreted, *a priori*, as being relatively small. After all, the estimate for the equivalent parameter in the labour market search-and-matching literature most often cited (Blanchard and Diamond, 1989) is $\eta = 0.6$. Following the argument in Cooley and Quadrini (1999a) however, a low value for η can be justified as a proxy for the endogenous effect on participation (in the case of the present paper, entry into entrepreneurship) that increased search efforts of banks imply.

We fix the remaining parameters of the matching function by requiring them to be consistent with the following steady state properties of the model: the probability that a searching entrepreneur finds a bank is 0.75, the probability that a loan officer finds an entrepreneur is 0.75 and the fraction of entrepreneurs that are active, ve, is equal to 0.95.[22] In future versions of the paper, we plan to appeal to surveys and studies of small business lending to justify better these values.

Finally, the parameter governing the sharing of the surplus in a bank-entrepreneur match (ξ) is such that the steady state spread between the (annualized) lending and deposit rates is around 2%.[23] The steady state equations then are used to back out the remaining parameters of the search technology as well as the exogenous rate of match separation.

The calibration of the household side of the model is standard. The disutility of leisure (represented by the function $l(N_t^w)$ has a fairly high curvature, so that the real-wage elasticity of labour supply is 3.0 in steady state. The utility parameters are such that a typical unskilled worker supplies 20% of his available hours to the labour market.

We compute a linear approximation to the solution of this system around its non-stochastic steady-state, with the help of the solution technique described in King and Watson (1998).

[22] These parameter values imply that the matching mechanism is very stable: the separation rate is lower than 5%. The general sense that one should get from our calibration of the lending relationships between entrepreneurs and banks is that they are fairly stable, compared to the volatile nature of firm-worker matches modelled elsewhere.

[23] This yields a value of around 0.1 (10%) for the share of the surplus going to a bank.

5. Results

5.1 Bargaining weights and persistence

Considering the calibration choices, as well as using the first order condition (14), allows us to rewrite equation (16) (defining the lending rate) as follows:

$$F(n_t^e, n_t^w)[(\frac{R_t^l}{R_t^d} - (1-\xi))\alpha_2 - \xi] + \frac{(R_t^d - 1)\, sc\, (1-\xi)}{\phi_t} = 0. \qquad (37)$$

This equation helps in understanding the mechanism by which the search-and-matching environment employed in the present paper can deliver some persistence to the liquidity effect following a monetary policy shock. A decrease in the tightness of the loan market from the point of view of the entrepreneurs (i.e. a decrease in the variable ϕ_t) increases, all things equal, the second term of (37). If one takes the value of R_t^d as mostly determined by the aggregate amount of liquidity available in the market for money, this means that the first term of (37) must decrease. One way for that decrease to occur is for the intermediation spread, $\frac{R_t^l}{R_t^d}$, to decrease. In effect, the decrease in the tightness variable ϕ_t signifies that it is easier for inactive entrepreneurs to enter the loan market and this increases their value function. In turn, the surplus enjoyed by active entrepreneurs (relative to a situation where they would be inactive) decreases. To compensate their existing clients for such a decrease in their surplus, banks have only the instrument of the lending rate.[24] A decrease in the intermediation spread is thus a natural outcome of this situation.

Consider now a liquidity injection from monetary authorities. This puts general downward pressure on the deposit rate R_t^d and upwards pressure on the scale of operations of existing clients. Banks increase their search efforts to find establish new relationships. This increase in the number of 'vacancies' (v_t^b), for a given number of inactive entrepreneurs $(1-v_t^e)$, initiates a decrease in the tightness variable ϕ_t.[25] The effect described above then creates downward pressure on the intermediation spread. In the subsequent periods, the number of active entrepreneurs has increased (as a result of the increase search efforts of banks) and thus the number of active entrepreneurs $(1-v_t^e)$ is now smaller than steady-

[24] Yet another way to interpret the situation is that the outside option of active entrepreneurs has improved and thus banks must ameliorate the terms of the lending contract with these clients in order for them not to terminate the relationship.

[25] Recall that the definition of the tightness variable is $\phi_t = \frac{1-v_t^e}{v_t^b}$

state. It is this effect that now continues to put downward pressure on ϕ_t and thus on the intermediation spread.

Further specialize (37) by assuming that all the surplus goes to banks ($\xi = 1$). The equation now states the following:

$$\frac{R_t^l}{R_t^d} = 1/\alpha_2.$$

In such a case, because banks get all the surplus from a match, they do not need to take the bargaining power of entrepreneurs in consideration and thus the intermediation spread does not react to loan market tightness. Further, as α_2 tends to 0, ever increasing rents are present in production and banks receive the totality of these rents by increasing without bound their intermediation spread. In contrast, assuming that $\xi = 0$ (entrepreneurs get all the surplus from a match) leads to this version of (37):

$$F(n_t^e, n_t^w)[\,(\frac{R_t^l}{R_t^d} - 1)\,\alpha_2\,] + \frac{(R_t^d - 1)sc}{\phi_t} = 0.$$

The effect described above is now maximized because a decrease in ϕ_t now increases by the same amount the second term on the equation.

5.2 Quantitative results

To test the quantitative importance of the effects mentioned above, Figure 1 and 2 show the impulse responses of some of the economic variables of interest following a liquidity injection. This liquidity injection, or monetary policy easing, represents a transitory increase in the rate of money growth from its steady-state level of 2% to a rate of 6%. In the graphs, the shock occurs in period 5. Figure 1 displays the responses of the real variables (production and matching market) while Figure 2 presents the responses of the nominal variables (interest rates, money and inflation variables).

Figure 1 shows that, following the increase in search efforts by banks, the probability that a searching entrepreneur finds a suitable loan officer increases on impact and remains high for several periods. Conversely, the probability that a searching loan officer finds a suitable entrepreneur decreases and only slowly comes back to control. This is the result of a congestion effect: the increased search efforts by all loan officers mean that the efforts of one are activities impede on the chance of success of another.

The output of a typical firm (not shown on the figure) increases at first but then decreases and converges from below. This is because in the initial period, the number of active entrepreneurs to lend to is limited and therefore banks choose to increase somewhat their lending to these few clients. In the following periods however, search has produced several new clients : see the bottom-right panel of the Figure. Banks now prefer to lend a bit less to each client so as to spread available liquidities to the maximum number of clients. The increased efficiency of the economy (brought about because spreading the same liquidities to an increased number of entrepreneurs allows the economy to escape the diminishing returns to scale) shows up in the responses of aggregate output and employment, which only increase significantly in the second period. Notice that our framework is thus capable of generating responses in output and employment that have a hump-shape component; such shape, not generated by the standard model, is thought of characterizing the empirical responses of variables to a monetary easing.

The bottom-left panel of the figure reports that the increased search activities of banks leads to a burst in the rate at which matches are created.[26] This burst in the match-creation rate is short lived and the rate actually converges from below in subsequent periods. But the initial effect is strong enough to lead to a persistent increase in the fraction of entrepreneurs that are active (have a lending relationship with a bank) and a persistent decrease in the loan market tightness (see the middle- and bottom-right panels of the figure). The rate at which matches are destroyed does not change because it is exogenous in the simplified model. The consequences of allowing this rate to be endogenous are explored in the following section.

Turning to the responses of the nominal variables in Figure 2, recall for a moment what these responses look like in the standard limited participation framework. The liquidity injection, deposited in the banking sector too late for households to change their deposit decisions, results in a relative abundance of liquidity in the banking sector. Banks

[26]The rate of match creation is defined as the number of new matches over the stock of existing matches. Writing that rate as jcr_t and using equations (17) and (21) leads to the following expression for that rate:

$$jcr_t = \frac{\chi_t(1 - v_t^e)\phi_t^{\eta-1}}{v_t^e}.$$

significantly reduce their lending rate to entice firms to increase their borrowing. Firms increase their hiring and thus their output. By the beginning of the next period however, households have corrected the initial imbalance and all nominal variables are essentially back to their pre-shock levels. Finally, the movements in the lending and deposit rates match one an other perfectly so that the intermediation spread remains constant.

In the present environment, the same aggregate imbalance of funds is present. However, banks allocate a portion of that surprise injection to the search for new clients, as can be seen in the bottom-right panel of Figure 2.[27] Because of the bargaining linking entrepreneurs and banks discussed above, not only the lending rate but also the intermediation spread declines on impact. The deposit rate also decreases on impact, although the amplitude of the response is smaller. Further, the rates and the spread only converge back progressively to their steady state levels. The persistent decrease in the intermediation spread shown in Figure 2 is also present in Cook (1999), where it arises because of the decreases in intermediation costs brought about by the initial increase in lending (through a learning-by-doing mechanism).[28] In contrast, this persistence is, in the present paper, related to the changes in the relative bargaining power of loan officers and entrepreneurs that the initial shock has created (recall the persistent decrease in the market tightness ϕ_t shown in Figure 1).

Finally, the fact that only part of the money injection is circulated in the economy in the impact period (the rest being held for search) helps explain the lagged response of inflation to the money injection: while inflation does increase somewhat in the impact period, its largest increase occurs in the period following the shock. Again, this is an aspect of our environment that generates responses much more consistent with the empirical evidence than the standard model (in which the responses of money and inflation would overlap almost perfectly) does.

To make the comparison clearer, Figure 3 presents the responses of a few variables in the standard limited participation model in the responses arising from our environment.

[27]In the language of Section 3.4, 'seed money' is equal to $\frac{S_t}{M_t}$, the funds that banks allocate to the search for new clients rather than making them available for lending to their existing ones.

[28]Cook (1999) also reports empirical evidence (arising from a VAR analysis) that suggests such responses in the intermediation spread are a feature of actual data.

The delayed, hump-shaped responses of output, employment and inflation are clearly visible when compared with those in the standard model. Further, the persistent decline in the intermediation spread arising from our set-up is in clear contrast with the constant spread present in the standard model.[29]

5.3 Other disturbances: shocks to technology and to the matching function

TO BE ADDED

6. Conclusion

This paper embedds elements of the search-and-matching litereature, applied to the relationship between borrowers and lenders, into the standard limited participation model. Specifically, it assumes that existing clients of banks operate a diminishing returns to scale technology, and second, that banks enter into contact with suitable new clients via a search-and-matching mechanism similar to that used to study labour markets (Pissarides, 1985). In such a framework, the initial liquidity injection that acoompagnies a moentary easing generates persistent changes in the relative bargaining weights of banks and entrepreneurs, affecting the sharing of net surpluses and thus the nominal interest rate that implements it. This effect is shown to add persistence to the decrease in nominal interest rates that follows monetary policy easings and to create hump-shaped responses in output and inflation.

In future work, we plan to experiment with a version of the model that would allow the rate at which bank-entrepreneur matches are terminated to endogenously respond both to the underlying quality of the match and business cycle[30]. The cyclical behaviour of job-match destruction has been shown to possess interesting cyclical properties and thus similar properties for bank-entrepreneur matches may enrich the analysis. Further, we would like to extend the model to one where households are risk-averse and physical capital accumulation takes places: while we do not expect the qualitative nature of our

[29] The intermediation spread is actually constant at a value of 0 in the standard model. To facilitate graphical comparisons, we have scaled it up.

[30] See for example, Cooley and Quadrini (1999a)

results to be modified by such an extension, it would allow our results to be more easily compared with the standard limited-participation literature that routinely includes these two features.

References

Andolfatto, D. 1996. "Business cycles and labour-market search." *American Economic Review* 86: 112-132.

Andolfatto, D. and P. Gomme. 2003. "Monetary Policy Regimes and Beliefs." *International Economic Review* 44: 1-30.

Andolfatto, D., S. Hendry, and K. Moran. 2002a. "Labour Markets, Liquidity, and Monetary Policy Regimes." Bank of Canada Working Paper No. 2002-32. Forthcoming, *Canadian Journal of Economics.*

——. 2002b. "Inflation Expectations and Learning about Monetary Policy." Bank of Canada Working Paper No. 2002-30.

Alvarez, F. and A. Atkeson. 1997. "Money and Exchange Rates in the Grossman-Weiss-Rotemberg Model." *Journal of Monetary Economics* 40: 619-640.

Becsi, Z., V. Li and P. Wang. 2000. "Financial matchmakers in Credit Markets with Heterogeneous Borrowers." Working Paper No. 2000-14, Federal Reserve Bank of Atlanta.

——. 2002. "Mismatch in Credit Markets." Manuscript, Vanderbilt University. February.

Berger, A.N. and G.F. Udell. 2002. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." *The Economic Journal* 112: F32-F53.

Bernanke, B., M. Gertler and S. Gilchrist. 1999. “The Financial Accelerator in a Quantitative Business Cycle Framework.” In Handbook of Macroeconomics, Volume1C, J. Taylor and M. Woodford editors, Amsterdam: Elsevier.

Blanchard, O. and P. Diamond. 1989. “The Beveridge Curve.” *Brookings Papers on Economic Activity* 1: 1–60.

Carlstrom, C. and T. Fuerst. 2001. “Monetary shocks, agency costs, and business cycles.” *Carnegie-Rochester Conference Series On Public Policy* 54(1): 1-27.

Christiano, L. J. 1991. “Modelling the Liquidity Effect of a Monetary Shock.” *Federal Reserve Bank of Minneapolis Quartely Review* 15(Winter): 3-34.

Chari, V.V., L. Christiano and M. Eichenbaum. 1995. “Inside Money, Outside Money and Short-Term Interest Rates.” *Journal of Money, Credit and Banking* 27(4) Pt. 2: 1354-1386.

Christiano, L. and M. Eichenbaum. 1998. “Modelling Money.” Working Paper 6371, National Bureau of Economic Research, January.

Christiano, L. and C. Gust. 1999. “Taylor Rules in a Limited Participation Model.” Working Paper 7017, National Bureau of Economic Research, March.

Cole, R.A. 1998. “The Importance of Relationships to the Availability of Credit.” *Journal of Banking and Finance* 22: 959-77.

Cooley, T. and V. Quadrini. 1999a. "A neo-classical model of the Phillips curve relation." *Journal of Monetary Economics* 44: 165-193.

Cooley, T. and V, Quadrini. 1999b. "Monetary Policy and the Financial Decisions of Firms." Manuscript, NYU University.

Cook, D. 1999. "The liquidity effect and money demand." *Journal of Monetary Economics* 43: 377-390.

Den Haan, W.J., G. Ramey and J. Watson. 1999. "Liquidity Flows and Fragility of Business Enterprises." University of California at San Diego Dicussion Paper No. 99-07R, December.

Engineer, M. and S. Shi. 1998. "Asymmetry, Imperfectly Transferable Utility, and the Role of Fiat Money in Improving Terms of Trade." *Journal of Monetary Economics* 41: 153-83.

Gertler, M. and S. Gilchrist. 1994. "Monetary Policy, Business Cycles, and the Behaviour of Small Manufacturing Firms." *The Quartely Journal of Economics* 109(2): 309-340.

King, R.G. and Watson, M.W. 1998. "The solution of linear difference systems under rational expectations." *International Economic Review*, 39(4): 1015-1026.

Koskela, E. and R. Stenbacka. 2001. "Equilibrium unemployment with credit and labour market imperfections." Bank of Finland Discussion paper 2001-05.

Merz, M. 1995. "Search in the labor market and the real business cycle." *Journal of Monetary Economics* 36(2): 269-300.

Petersen, M.A. and R.G. Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data ." *Journal of Finance* Vol. XLIX No. 1: 3-37.

Petrongolo, B. and C.A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* Vol. XXXIX (June): 390-431.

Pissarides, C. A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies and Real Wages." *American Economic Review* 75(4): 676-690.

Shimer, R. 2003. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies: Evidence and Theory." NBER Working Paper No. 9536 .

Wasmer, E. and P. Weil. 2000. "Credit Markets and Unemployment in the Short Run and in the Long Run." *International Economic Journal* 15(1): 1-19.

Figure 1: Responses to a monetary easing, real variables

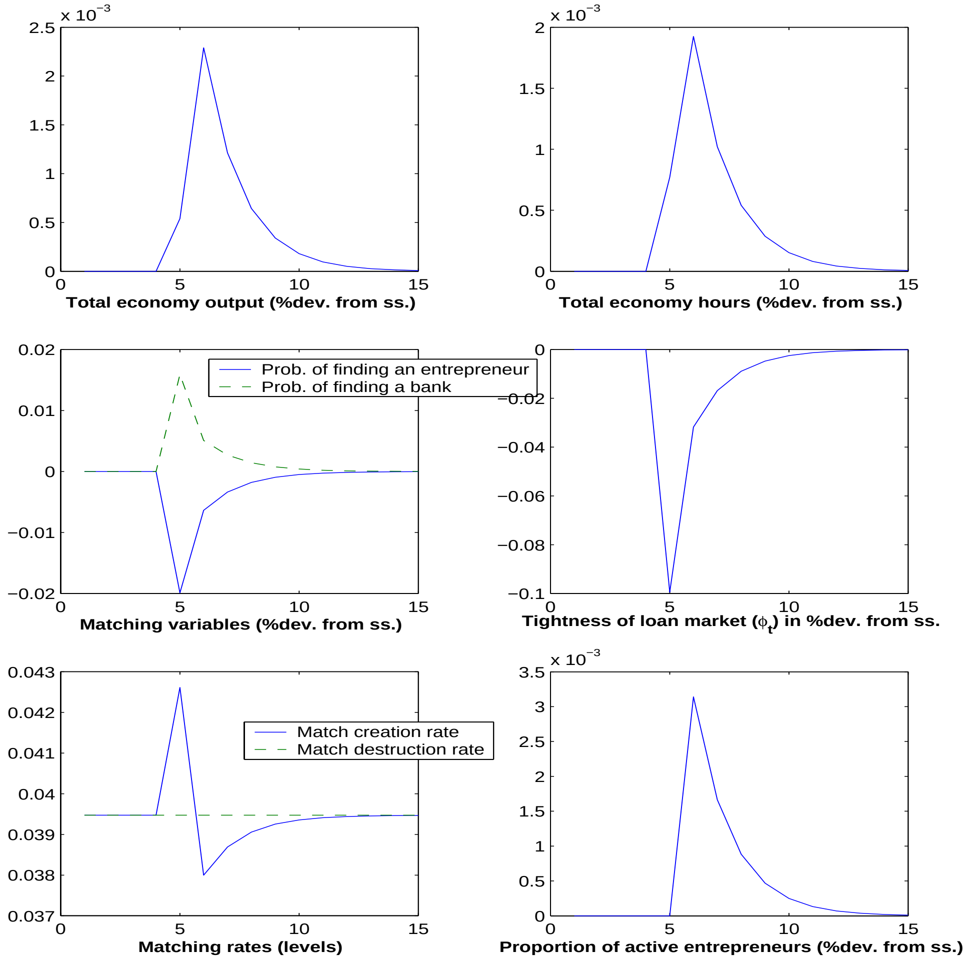

Figure 2: Responses to a monetary easing, nominal variables

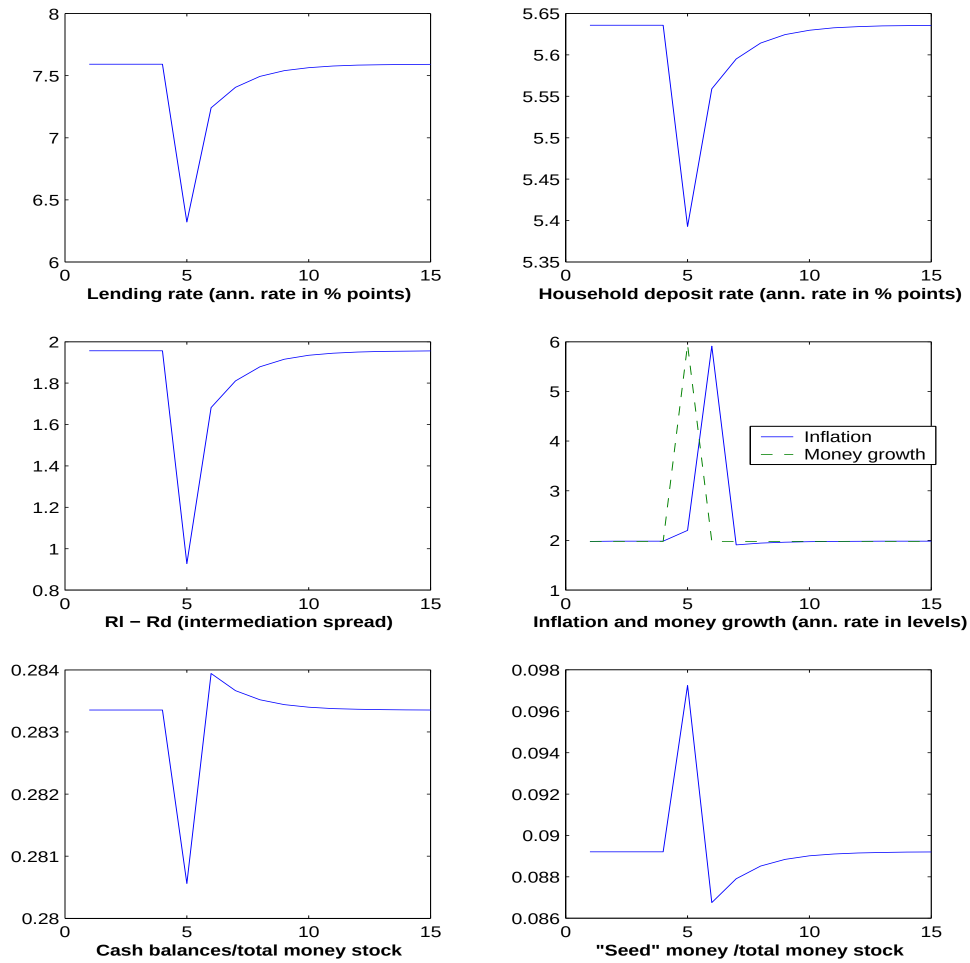

Figure 3: Responses to a monetary easing: comparison with the standard Limited Participation Model