Equilibrium Unemployment, Job Flows and Inflation Dynamics*

Antonella Trigari†
IGIER, Università Bocconi

First version: November 2003
This version: May 2008

Abstract

In order to explain the joint fluctuations of output, inflation and the labor market, this paper develops and estimates a general equilibrium model that integrates a theory of equilibrium unemployment into a monetary model with nominal price rigidities. The estimated model accounts for the responses of employment, hours per worker, job creation and job destruction to a monetary policy shock. Moreover, search frictions in the labor market generate a lower elasticity of marginal costs with respect to output. This helps to explain the sluggishness of inflation and the persistence of output that are observed in the data.

Keywords: DSGE Models, Sticky Prices, Labor-Market Search, Monetary Policy Shocks

*I wish to thank Mark Gertler, Ricardo Lagos, and Vincenzo Quadrini for their guidance. I am also indebted to Jean Boivin, Fabio Canova, Marc Giannoni, Thomas Sargent, Ulf Soderstrom, Gianluca Violante, and Raf Wouters for helpful suggestions and discussions. Many thanks also to an anonymous referee. This paper benefitted from comments by seminar participants at the 2003 meetings of the Society for Economics Dynamics, UCSB, New York University, the Bank of England, CREI, the Stockholm School of Economics, the Federal Reserve Bank of New York, IGIER, the Federal Reserve Bank of Boston, Indiana University, UAB, the Federal Reserve Bank of St. Louis, Tufts University, the Federal Reserve Bank of Kansas City, the Federal Reserve Board, Tilburg University, the Ente Einaudi, the European Central Bank and George Washington University. Remaining errors are my responsibility.

†Correspondence: IGIER - Università Bocconi, Via Roentgen, 20136 Milano - Italy. E-mail: antonella.trigari@unibocconi.it. Home page: www.igier.unibocconi.it/trigari. Tel. +39.02.58363040. Fax +39.02.58363302.

1 Introduction

A classic challenge that macroeconomists face is to explain the cyclical fluctuations of output, unemployment and inflation. Recently, New Keynesian (NK) business cycle models have made important advances in explaining the links between monetary policy and the joint dynamics of output and inflation.[1] However, the standard NK model abstracts from unemployment as it assumes a neoclassical labor market in which individuals vary the hours that they work, but the number of people working never changes. This of course implies that the model cannot account for evidence regarding the effects of aggregate shocks, in particular monetary policy shocks, on unemployment dynamics. Moreover, when accounting for the joint response of output and inflation to monetary policy shocks, the standard NK model has a great difficulty in replicating the sluggish response of inflation together with the large and persistent response of output. One key reason for this difficulty is that the model has the labor input adjusting along the intensive margin, which makes real wages very responsive over the cycle unless an implausibly high labor supply elasticity is assumed. This in turn induces firms setting prices as a markup over marginal cost to make large price adjustment and causes inflation in the model to fluctuate more than evidence suggests. Based on these and related considerations, several recent papers have argued that labor market frictions are crucial to understanding business cycle fluctuations, as well as the effects of monetary policy shocks and the design of monetary policies.[2] The search and matching model, along the lines of the work of Mortensen and Pissarides (1994), is a natural way of thinking about these frictions.

In this paper I develop a dynamic general equilibrium model that integrates labor market search and endogenous job destruction into an otherwise standard NK model with nominal price rigidities.[3] I show that introducing search and matching frictions modifies the nature of real marginal cost faced by firms in a way that lowers the elasticity of marginal cost with respect to output and thus helps to account for the observed inertia in inflation and persistence in output. To gain some intuition, first note that in the model changes in output can be obtained through either changes in the number of hours worked, the intensive margin, or changes in the number of employed people, the extensive margin. With demand-constrained firms, the two margins will adjust to meet demand so that their marginal costs are equalized. Now, although hours worked by each employed worker are chosen through bargaining, the marginal cost of hours is determined by the workers' marginal disutility from supplying labor, much the same way as in a neoclassical labor market. However, while in a neoclassical framework all variation of the labor input occurs at the intensive margin, which is very costly when the labor supply elasticity is not implausibly high, with equilibrium unemployment

[1] See Galì (2003) for a survey.

[2] Among these see Galì, Gertler and Lopez-Salido (2001), Smets and Wouters (2003), Christiano, Eichenbaum and Evans (2005), Levin, Onatski, Williams and Williams (2005).

[3] Early work by Merz (1995), Andolfatto (1996), den Haan, Ramey and Watson (2000) and Hairault (2002) has considered search and matching in a real business cycle model. Cooley and Quadrini (1999) integrate a model of equilibrium unemployment with a limited participation model of money.

firms can change employment at the extensive margin. The cost of adjusting output through changes in the number of employed workers in a framework with hiring frictions and long-term employment relationships has two components. The marginal cost increases with the flow cost of having one additional worker employed but decreases with the expected future payoff from continuing the relationship in the following periods. These dynamic considerations make the extensive margin relatively less costly and induce firms to adjust the labor input mostly at the extensive margin, as we observe in the data.

After developing the theoretical model, I estimate a set of the structural parameters that characterize the dynamics of the labor market and on which there is little or no independent evidence. I follow the limited information estimation strategy adopted in Rotemberg and Woodford (1997) and others. Specifically, the structural parameters are chosen so that the impulse responses to a monetary policy shock of a set of endogenous variables in the model match as closely as possible the responses estimated using a Vector Autoregressive (VAR) methodology. While the minimum distance estimation strategy is widely adopted in the literature on dynamic general equilibrium models with money, no other study, to the best of my knowledge, has used it to estimate the parameters of a labor market characterized by matching frictions and endogenous job destruction.

In order to evaluate the model I proceed in two steps. First, I assess the contribution of labor market frictions in shaping the joint dynamics of output and inflation by comparing the predictions of the model developed in this paper with those of a NK model that does not have search and matching frictions but keeps all other features the same. I show that in the model with labor market search the response of inflation is significantly less volatile and the response of output considerably more persistent than in the baseline NK model. Second, I evaluate the ability of the model to account quantitatively for the response of the U.S. economy to a monetary policy shock. The estimated model does a good job in replicating the observed responses of output, inflation and the labor market (specifically employment, hours per worker, job creation and job destruction).

Independent work by Walsh (2005) also studies the interaction between price rigidities and labor-market search. There are three main differences with this paper. First, his work considers only the extensive margin, while I consider the intensive as well as the extensive margin. This allows me to explain the dynamics of hours per worker over the cycle as well as the dynamics of employment and to clarify how the interaction between the two margins shapes real marginal costs. Second, rather than taking a stand on all the possible sources of fluctuations in the economy, I evaluate the empirical performance of the model based on its ability to match conditional second moments, i.e., second moment conditional on a particular source of fluctuations.[4] Third, I provide estimates of a set of the structural parameters that characterize a labor market with search and matching frictions.

The remainder of the paper is organized as follows: Section 2 presents the evidence related to the response of output, inflation and the labor market to a monetary policy shock; Section 3

[4]The advantages of this evaluation criterion are presented in Galí (1999).

describes the model; Section 4 analyses the determination of real marginal costs with search and matching frictions; Section 5 brings the model to the data and discusses the estimation; Section 6 presents the results; Section 7 discusses the alternative approach to modeling labor market frictions, i.e., nominal wage rigidities; and Section 8 concludes.

2 Evidence: output, inflation and the labor market

In this section I describe a set of stylized facts related to the behavior of output, inflation and a set of labor market variables in face of a monetary policy shock. The evidence regarding the response of the labor market to a monetary policy shock is new in the literature. More specifically, I use a VAR methodology to estimate the dynamic response of the variables of interest to an identified exogenous monetary policy shock. The short-term nominal interest rate is taken to be the instrument of monetary policy and the identification strategy is based on the recursiveness assumption as in Christiano, Eichenbaum and Evans (2000).[5]

The variables included in the analysis are measures of output, inflation and the nominal interest rate, to which I add four labor market variables. The labor market variables that I include are measures of employment, average hours per worker, the job creation rate and the job destruction rate. I include four lagged values of all variables in the VAR. Estimates are based on quarterly U.S. data from 1972:2 to 1993:4.[6]

The series for the nominal interest rate is the federal funds rate, annualized and averaged over the quarter. The series for output is the log of quarterly real GDP and the series for inflation is the annualized rate of change of the GDP deflator between two consecutive quarters. The series for employment is the log of total employees in nonfarm establishments. The series for average hours per worker is constructed by subtracting the previous variable from the log of total employee-hours in nonagricultural establishments. Finally, the series for job creation and job destruction are taken from Davis, Haltiwanger, and Schuh's "Job Creation and Destruction" database. They are, respectively, the log of the quarterly job creation rate for both startups and continuing establishments in the manufacturing sector and the log of the quarterly job destruction rate for both shutdowns and continuing establishments in the manufacturing sector.

Figure 1 reports the responses over time of output, inflation and the federal funds rate to a one percent increase in the federal funds rate and Figure 2 the responses of employment, average hours per worker, the job creation rate and the job destruction rate to the same shock. The solid lines display the point estimates of the coefficients. The dashed lines are two standard deviation confidence intervals. The impulse response functions of inflation and the federal funds rate are reported in percentage points. The other impulse responses are reported in percentage deviations from each variable's unconditional mean. The horizontal axes indicate quarters.

[5] The details of the identification strategy are described in a technical appendix available at the author's website.
[6] The choice of the sample period is explained by the availability of the data on job creation and job destruction.

The results suggested by Figure 1 are standard in the VAR literature on monetary policy. After a contractionary monetary policy shock there is a large hump-shaped fall in output accompanied by a sluggish persistent decrease in inflation. The peak fall in output is about 0.4 percent and that of annualized inflation about 0.3 percent. Existing optimizing monetary general equilibrium models have shown difficulties in explaining this joint dynamic behavior of output and inflation. In their baseline version, unless a large value of the labor supply elasticity is assumed, they predict a much larger response of inflation.

Figure 2, instead, presents some new results about the response of the labor market to a monetary policy shock. First, the labor input adjusts along both the extensive and the intensive margin. As a consequence of the tightening in monetary policy, both employment and hours per worker fall. However, while the fall in employment is large and persistent, there is only a small transitory decrease in hours per worker. Therefore, the labor input shows a significantly different cyclical behavior at the extensive and the intensive margin. Second, the response of employment is explained by variations at both the job creation and the job destruction margin. The monetary contraction causes a fall in job creation and an increase in job destruction. The decrease in job creation is transitory with a peak response of about 3.4 percent, while the increase in job destruction is larger and more persistent with a peak response of about 4.5 percent.

Finally, the tightening in monetary policy has a significant effect on output, employment, hours per worker, job creation and job destruction only after two quarters.[7] In order to reproduce this feature of U.S. data, following Rotemberg and Woodford (1997) and others, I will introduce in the model informational lags in the decisions to consume, set prices, post vacancies and endogenously severe a match.

3 The model

The proposed model with nominal price rigidities and search and matching in the labor market has four sectors: households, intermediate goods firms, retail firms and a monetary authority.

3.1 Households

Each household is thought of as a large extended family which contains a continuum of members with names on the unit interval. In equilibrium, some members will be unemployed while some others will be working for intermediate goods firms. Each member has the following period utility function:

$$u(c_t, c_{t-1}) - g\left(h_t, a_t\right), \tag{1}$$

with

$$u(c_t, c_{t-1}) = \log\left(c_t - ec_{t-1}\right), \tag{2}$$

[7] At the same time, the response of inflation presents the well known "price puzzle".

and

$$g\left(h_t, a_t\right) = \kappa_h \frac{h_t^{1+\phi}}{1+\phi} + \chi_t a_t, \tag{3}$$

where c_t is consumption of a final good, h_t is the hours of work, a_t is a shock to the disutility from working and χ_t is an indicator function taking the value of one if the individual is employed and zero if unemployed. When $e > 0$, the model allows for habit formation in consumption.[8] The preference shock a_t is idiosyncratic to the individual and is assumed to be independently and identically distributed across individuals and times with cumulative distribution function $F\left(a_t\right)$. The cumulative distribution function $F\left(a_t\right)$ is assumed to be lognormal with parameters μ_a and σ_a. Assuming that the idiosyncratic shock enters additively avoids the problem of excessive variation in hours worked across individuals. In particular, since individuals are identical in all aspects other than the preference shock, it will be the case that they all work the same number of hours.[9]

The representative household maximizes lifetime utility:

$$E_t \sum_{s=0}^{\infty} \beta^s \left[u(c_{t+s}, c_{t+s-1}) - G_{t+s}\right], \tag{4}$$

where $\beta \in (0,1)$ is the intertemporal discount factor and G_t denotes the household's disutility from supplying hours of work, i.e., the sum of the disutilities of employed members. The representative household does not choose hours of work: these are determined through decentralized bargaining between firms and workers, as we discuss below.[10]

Households own all firms in the economy and face each period the budget constraint

$$c_t + \frac{B_t}{p_t r_t^n} = d_t + \frac{B_{t-1}}{p_t}, \tag{5}$$

where p_t is the aggregate price level, B_t is holdings of a nominal one-period bond and r_t^n is the gross nominal interest rate on this bond. The variable d_t is the household's income in period t.[11]

The representative household chooses consumption and asset holdings to maximize (4) subject to (5), which yields

$$\lambda_t = \beta E_t \left(r_t \lambda_{t+1}\right), \tag{6}$$

where λ_t is the marginal utility of consumption and r_t is the gross real interest rate.

[8] See Section 6 for a discussion on the role of habit persistence in the model.

[9] The literature typically assumes multiplicative productivity shocks. See Cooley and Quadrini (1999) for an alternative example of additive idiosyncratic shocks.

[10] I follow Merz (1995) and others in assuming that household members perfectly insure each other against fluctuations in consumption. Since consumption is equalized across members, I use the same notation for consumption of the representative household and consumption of each member.

[11] Household income is wages earned by employed members, benefits earned by unemployed members and the share of aggregate profits from firms, net of a government lump-sum tax used to finance unemployment benefits.

3.2 Firms and the labor market

The model is characterized by two main building blocks: nominal rigidities in price setting and search and matching frictions in the labor market. One complication is that when firms set prices on a staggered basis the job creation and destruction decisions become intractable. To avoid this problem I distinguish between two types of firms: retailers and intermediate good firms, or simply firms.[12,13] Firms produce intermediate goods in competitive markets using labor as their only input, and then sell their output to retailers who are monopolistically competitive. Retailers sell final goods to the households. Then, I assume that price rigidities arise at the retail level, while search frictions occur in the intermediate good sector. In this section I describe the problem of intermediate goods firms.

3.2.1 Matching market and production

In order to match with a worker, firms must actively search for workers in the unemployment pool. This idea is formalized by assuming that firms post vacancies. On the other hand, unemployed workers must look for firms. I assume that all unemployed workers search passively for jobs.

Each firm has a single job that can either be filled or vacant and searching for a worker. Workers can be either employed or unemployed and searching for a job.[14]

Vacancies, v_t, are matched to workers seeking for a job, u_t, according to the CRS matching function

$$m_t = \sigma_m u_t^{\sigma} v_t^{1-\sigma}, \tag{7}$$

where σ_m is a scale parameter reflecting the efficiency of the matching process.

The probability that any open vacancy is matched with a searching worker at date t is denoted with q_t and is given by

$$q_t = \frac{m_t}{v_t}. \tag{8}$$

Similarly, the probability that any worker looking for a job is matched with an open vacancy at time t is denoted with s_t and is given by

$$s_t = \frac{m_t}{u_t}. \tag{9}$$

If the search process is successful, the firm operates a production function $f(h_t) = h_t$, where h_t is the time spent working at date t. Employment relationships might be severed for exogenous reasons at the beginning of any given period. I denote with ρ^x the probability of exogenous

[12] This modelling device was first used by Bernanke, Gertler and Gilchrist (1999) in their study of the financial accelerator mechanism. An alternative is to assume quadratic price adjustment costs, as in Krause and Lubik (2007).

[13] On top of retailers and intermediate good firms, final good firms combine individual retail goods into a final good using a Dixit-Stiglitz aggregator (see Section 3.3). Since this modelling assumption is standard in the NK literature (see Ireland, 1997, for an early example), I do not describe the problem of final good producers in much detail.

[14] I abstract from workers' labor force participation decisions.

separation. Furthermore, a matched pair may chose to separate endogenously. If the realization of the idiosyncratic preference disturbance a_t is above a certain threshold, which I denote $\underline{a}_t$, a firm and a worker discontinue their relationship. The probability of endogenous separation is $\rho_t^n = \Pr(a_t > \underline{a}_t) = 1 - F(\underline{a}_t)$ and the overall separation rate is $\rho_t = \rho^x + (1-\rho^x)\rho_t^n$. If either exogenous or endogenous separation occurs, production does not take place.

Employment evolves according to the dynamic equation

$$n_t = \left(1 - \rho_{t-1}\right) n_{t-1} + m_{t-1}, \tag{10}$$

which simply says that the number of matched workers at the beginning of period t, n_t, is given by the fraction of matches in $t-1$ that survives to the next period, $\left(1-\rho_{t-1}\right) n_{t-1}$, plus the newly-formed matches, m_{t-1}.

The labor force being normalized to one, the number of unemployed workers at the beginning of period t is $1 - n_t$. This is different from the number of searching workers in period t, u_t, which is given by

$$u_t = 1 - (1 - \rho_t) n_t, \tag{11}$$

since some of the employed workers discontinue their match and search for a new job in the same period.

3.2.2 Bellman equations

To make the exposition of the following sections easier, I describe here the Bellman equations that characterize the problem of firms and workers.

Denote with J_t the value of a job for a firm at date t measured in terms of current consumption of the final good. This is given by

$$J_t(a_t) = x_t f(h_t) - w_t(a_t) h_t + E_t \beta_{t,t+1} \left(1 - \rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} J_{t+1}(a_{t+1}) \frac{dF(a_{t+1})}{F(\underline{a}_{t+1})}, \tag{12}$$

where x_t and w_t denote, respectively, the relative price of the intermediate good and the hourly wage rate at date t. Note that the hourly wage rate depends on the realization of the preference shock. The value of the job is the current profits $x_t f(h_t) - w_t(a_t) h_t$ plus the continuation value. Next period, with probability $1 - \rho_{t+1}$ the match is not severed. In this event the firm obtains the future expected value of a job, where the expected value is conditional on having the preference shock a_{t+1} below the separation threshold $\underline{a}_{t+1}$. With probability ρ_{t+1}, instead, the match is discontinued in $t+1$ and the firm obtains a future value equal to zero. Finally, the expected future value of the job is discounted by the factor $\beta_{t,t+1}$, where $\beta_{t,t+s} = \beta^s \lambda_{t+s}/\lambda_t$.[15]

[15] The use of this discount factor effectively evaluates profits in terms of the values attached to them by the households, who ultimately own firms.

Denote with V_t the value of an open vacancy for a firm at date t expressed in terms of current consumption. Letting κ be the utility cost of keeping a vacancy open, V_t can be written as

$$V_t = -\frac{\kappa}{\lambda_t} + E_t \beta_{t,t+1} \left[q_t \left(1 - \rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} J_{t+1}\left(a_{t+1}\right) \frac{dF(a_{t+1})}{F\left(\underline{a}_{t+1}\right)} + \left(1 - q_t\right) V_{t+1} \right], \tag{13}$$

where κ/λ_t is the utility cost expressed in terms of current consumption.

Denote now with W_t and U_t, respectively, the employment and the unemployment value for a worker at date t expressed in terms of current consumption.[16] The value of employment W_t can be written as

$$W_t\left(a_t\right) = w_t\left(a_t\right) h_t - \frac{g\left(h_t, a_t\right)}{\lambda_t} + E_t \beta_{t,t+1} \left[\left(1 - \rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} \left(W_{t+1}\left(a_{t+1}\right) - U_{t+1}\right) \frac{dF(a_{t+1})}{F\left(\underline{a}_{t+1}\right)} + U_{t+1} \right], \tag{14}$$

where $g\left(h_t, a_t\right)/\lambda_t$ is the disutility from supplying hours expressed in terms of current consumption.

Finally, the value of unemployment U_t is given by

$$U_t = b + E_t \beta_{t,t+1} \left[s_t \left(1 - \rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} \left(W_{t+1}\left(a_{t+1}\right) - U_{t+1}\right) \frac{dF(a_{t+1})}{F\left(\underline{a}_{t+1}\right)} + U_{t+1} \right], \tag{15}$$

where b is the flow value of being unemployed, taken to be unemployment benefits.

3.2.3 Vacancy posting

As long as the value of a vacancy is greater than zero, firms open new vacancies. In equilibrium, free entry ensures that $V_t = 0$ at any time t. This yields the vacancy posting condition

$$\frac{\kappa}{\lambda_t q_t} = E_t \beta_{t,t+1} \left(1 - \rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} J_{t+1}\left(a_{t+1}\right) \frac{dF(a_{t+1})}{F\left(\underline{a}_{t+1}\right)}, \tag{16}$$

which equates the expected cost of hiring a worker to the expected value of a match.

Equation (16) implies that, holding constant λ_t, a decrease in the sum of expected future profits must be associated with an increase in q_t. Given the specification of the matching function, this requires either a decrease in the number of vacancies posted, v_t, or an increase in the number of searching workers, u_t. The number of workers searching for a job may change on impact due to changes in the job destruction rate ρ_t. However, unless u_t increases significantly, the increase in q_t will be associated with a fall in v_t. Monetary policy shocks will affect the rate at which vacancies are posted and, consequently, employment through this mechanism.

Specifically, a persistent raise in the nominal interest rate, which results in an increase in the real interest rate due to price rigidities, reduces current and future aggregate demand. Since

[16] Because there is perfect income insurance it is not straightforward to define these values. A technical appendix available at the author's website shows how the worker surplus $W_t - U_t$ can be derived from the household's problem.

monopolistically competitive retailers produce to meet demand, this diminishes their current and future demand for intermediate goods, which they use as inputs. The resulting persistent decrease in the relative price of intermediate goods x_t leads to a fall in firms' expected future profits, decreasing vacancies today and employment next period.

Finally, equation (16) can be rearranged to a first-order difference equation in q_t:

$$\frac{\kappa}{\lambda_t q_t} = E_t \beta_{t,t+1} \left(1 - \rho_{t+1}\right) \left(x_{t+1} h_{t+1} - w_{t+1} h_{t+1} + \frac{\kappa}{\lambda_{t+1} q_{t+1}} \right), \tag{17}$$

where w_t is the aggregate wage:

$$w_t = \int_0^{\underline{a}_t} w_t \left(a_t\right) \tfrac{dF(a_t)}{F(\underline{a}_t)}. \tag{18}$$

3.2.4 Bargaining over wages and hours

Bargaining takes place along two dimensions, the real wage and the hours of work. I assume Nash bargaining. That is, the firm and the worker choose the wage w_t and the hours of work h_t to maximize the Nash product

$$\left(W_t \left(a_t\right) - U_t\right)^{\eta} \left(J_t \left(a_t\right) - V_t\right)^{1-\eta}, \tag{19}$$

where the first term in brackets is the worker's surplus, the second is the firm's surplus, and $\eta \in (0, 1)$ reflects the parties' relative bargaining power.

Because the firm and the worker bargain simultaneously about wages and hours, the outcome is (privately) efficient and the wage does not play an allocational role for hours.[17] The Nash bargaining model, in effect, is equivalent to one where hours are chosen to maximize the joint surplus of the match, while the wage is set to split that surplus according to the parameter η.[18]

The wage chosen by the match satisfies the optimality condition

$$\eta J_t \left(a_t\right) = \left(1 - \eta\right) \left(W_t \left(a_t\right) - U_t\right). \tag{20}$$

Substituting the expressions for the worker and firm surpluses, using also the vacancy posting condition (16), gives the wage equation

$$w_t \left(a_t\right) h_t = \eta \left(x_t f(h_t) + \frac{\kappa}{\lambda_t} \theta_t \right) + \left(1 - \eta\right) \left(\frac{g(h_t, a_t)}{\lambda_t} + b \right), \tag{21}$$

where $\theta_t = v_t/u_t$ denotes labor market tightness. The wage shares costs and benefits from the activity of the match according to the parameter η. The worker is rewarded for a fraction η of the

[17] The outcome predicted by the Nash bargaining model is generally *not* efficient from the viewpoint of society as a whole (Hosios, 1990).

[18] Trigari (2004) develops an alternative bargaining model where firms have the right-to-manage hours, implying that the choice of hours is not privately efficient from the perspective of each firm-worker match. The model delivers interesting implications for the relation between wages and marginal cost.

firm's revenues and savings of hiring costs and compensated for a fraction $1-\eta$ of the disutility he suffers from supplying hours of work and the foregone unemployment benefits.

The hours of work, h_t, chosen by the match satisfy the optimality condition

$$x_t f_h(h_t) = \frac{g_h(h_t, a_t)}{\lambda_t}, \tag{22}$$

where the value of the marginal product of labor is equated to the marginal rate of substitution between consumption and leisure. Thus, the first order condition determining hours worked is exactly the same as in a competitive labor market. This happens because the correct measure of the cost of hours to the firm is the marginal rate of substitution, rather than the wage. In other words, the wage is not allocational for hours. Finally, because the preference shock a_t is additive, optimal hours do not depend on its realization.

3.2.5 Endogenous separation

Let $S_t(a_t) = J_t(a_t) + W_t(a_t) - U_t$ be the joint surplus from the match. Using (12), (14) and (15), together with (20), yields the following expression for $S_t(a_t)$:

$$S_t(a_t) = x_t f(h_t) - \frac{g(h_t, a_t)}{\lambda_t} - b + (1-\eta s_t) E_t \beta_{t,t+1} \left(1-\rho_{t+1}\right) \int_0^{\underline{a}_{t+1}} S_{t+1}(a_{t+1}) \frac{dF(a_{t+1})}{F(\underline{a}_{t+1})}. \tag{23}$$

The total surplus equals current revenues net of the disutility from supplying hours and the foregone unemployment benefit, plus the continuation value. The latter gives the current expected value of future joint payoffs from continuing the relationship in the following periods.

A successful match is endogenously discontinued whenever the realization of the preference shock makes the value of the joint surplus equal to zero or negative. Then, the condition that implicitly defines the threshold value $\underline{a}_t$ is $S_t(\underline{a}_t) = 0$. Using equations (16) and (20) in (23), and rearranging yields

$$x_t f(h_t) - \frac{g(h_t, \underline{a}_t)}{\lambda_t} - b + \frac{1-\eta s_t}{1-\eta} \frac{\kappa}{\lambda_t q_t} = 0. \tag{24}$$

Equation (24) implies that a fall in the expected future joint payoffs from continuing the relationship, $\frac{1-\eta s_t}{1-\eta}\frac{\kappa}{\lambda_t q_t}$, must be associated with an increase in the current joint payoffs evaluated at the threshold value, $x_t f(h_t) - g(h_t, \underline{a}_t)/\lambda_t - b$. If the decrease in the expected future payoffs is caused by a persistent contractionary aggregate shock, current payoffs at any given realization of the preference shock a_t are falling as well. In this case, the increase in current payoffs can only be obtained through a decrease in $\underline{a}_t$. Monetary policy shocks will affect the separation decision of firms and workers and, consequently, employment through the above mechanism. A persistent increase in the nominal interest rate reduces current and future expected payoffs at any given realization of a_t. This, in turn, decreases the value of a_t above which the firm and the worker decide to separate. A lower threshold $\underline{a}_t$ raises the current separation rate ρ_t on impact and decreases the number of workers producing within the same period.

3.2.6 Job creation, job destruction and employment

I define labor market flows following den Haan, Ramey and Watson (2000). They begin with the observation that flows of workers out of employment relationships are larger than flows of jobs out of firms. This implies that a fraction of the firms experiencing separations from workers must attempt to refill the jobs left vacant and be successful at doing it within the same period. To take this observation into account, they assume that firms experiencing exogenous separations immediately repost the resulting vacancies, while firms experiencing endogenous separations do not. This implies that $\rho^x n_t$ separations are reposted and $q_t \rho^x n_t$ separations are refilled within the same period. Finally, they assume that a job is neither created or destroyed by a firm that both looses and gains a worker in the same period.

Job creation, then, is defined to be equal to the number of newly-created matches net of the number of matches serving to refill the reposted vacancies. The job creation rate is given by

$$jc_t = \frac{m_t}{n_t} - q_t \rho^x. \tag{25}$$

Job destruction, in turn, is defined as the total number of separations net of the number of separations that are reposted and successfully refilled. The job destruction rate is given by

$$jd_t = \rho_t - q_t \rho^x. \tag{26}$$

Employment growth, finally, can be written as

$$\frac{n_{t+1} - n_t}{n_t} = jc_t - jd_t. \tag{27}$$

3.3 Retailers and price setting

There is a continuum of monopolistically competitive retailers indexed by i on the unit interval. Retailers do nothing other than buy intermediate goods from firms, differentiate them with a technology that transforms one unit of intermediate goods into one unit of retail goods, and re-sell them to the households. Note that the relative price of intermediate goods, x_t, coincides with the real marginal cost faced by the retailers.

Let y_{it} be the quantity of output sold by retailer i. Final goods, denoted with y_t, are the following composite of individual retail goods:

$$y_t = \left[\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{28}$$

where $\varepsilon > 1$ is the elasticity of substitution.

As in Calvo (1983), in any given period each retailer can reset its price with a fixed probability $1-\psi$. Moreover, I follow Galí and Gertler (1999) by assuming that there are two types of retailers that differ in the way they reset prices. A fraction $1-\omega$ of the retailers, which are referred to as "forward-looking", set prices optimally, given the restriction on the frequency with which they can adjust their price. The solution to their problem gives:

$$p_t^f = \mu E_t \sum_{s=0}^{\infty} \omega_{t,t+s} x_{t+s}^n, \tag{29}$$

where $\mu = \frac{\varepsilon}{\varepsilon - 1}$ is the flexible-price markup and $x_t^n = p_t x_t$ is the nominal marginal cost at date t. Expected marginal costs are weighted with the expected discounted revenue share ω.[19]

The remaining fraction ω of the retailers, which are referred to as "backward-looking", instead obey the following rule of thumb:

$$p_t^b = \pi_{t-1} \overline{p}_{t-1}, \tag{30}$$

where π_t is the gross inflation rate in t and $\overline{p}_t$ is the average of the newly reset prices at date t.[20]

Finally, the model is closed by imposing the economy-wide resource constraint

$$c_t = y_t, \tag{31}$$

and the market clearing condition in the intermediate good sector

$$y_t = \overline{n}_t h_t, \tag{32}$$

where y_t is aggregate demand, $\overline{n}_t = n_t (1 - \rho_t)$ is the number of firms producing in t and h_t is each firm's level of production.

3.4 Monetary authority

The monetary authority conducts monetary policy using the short-term nominal interest rate as the policy instrument. The gross nominal interest rate r_t^n follows the Taylor-type rule

$$\frac{r_t^n}{r^n} = \left(\frac{r_{t-1}^n}{r^n} \right)^{\rho_m} \left(\frac{E_t \pi_{t+1}}{\pi} \right)^{\gamma_\pi (1-\rho_m)} \left(\frac{y_t}{y} \right)^{\gamma_y (1-\rho_m)} e^{\varepsilon_t^m}, \tag{33}$$

where ρ_m measures the degree of interest rate smoothing, γ_π and γ_y are the response coefficients to inflation and output and variables without a time subscript denote steady state values. Finally, ε_t^m is an i.i.d. monetary policy shock.

[19] Specifically, $\omega_{t,t+s} = \frac{\psi^s \beta_{t,t+s} R_{it,t+s}}{E_t \sum_{k=0}^{\infty} \psi^k \beta_{t,t+k} R_{it,t+k}}$, where $R_{it,t+s}$ are revenues from good i at time $t+s$ conditional on the price set at date t.

[20] Having backward-looking retailers introduces a lagged inflation term in the Phillips curve and helps explaining inflation inertia. This assumption is essentially equivalent to the alternative one made in the literature that non-reoptimized prices are indexed to lagged inflation, as in Christiano, Eichenbaum and Evans (2005).

3.5 Limited information decisions and model dynamics

Following Rotemberg and Woodford (1997) and others, I assume that households must choose their consumption level at date t with the information set available at date $t-2$. In addition, I make a similar assumption for the decisions to set prices, post vacancies and separate. These assumptions are consistent with the identifying restriction imposed in the VAR considered in Section 2, according to which the variables in the information set of the central bank do not respond contemporaneously to a monetary policy shock.[21]

The model dynamics are obtained by taking a loglinear approximation around a deterministic steady state with zero inflation. The model nests the standard NK model with a neoclassical labor market. The NK model dynamics can be obtained by assuming that the rates of job creation and job destruction are constant at their steady state values. A technical appendix available at the author's web site presents the complete loglinear model.

4 Real marginal costs with search and matching frictions

Inflation dynamics in the model are determined by the behavior of real marginal cost, according to a conventional NK Phillips curve. What changes relative to the baseline framework with a neoclassical labor market is the behavior of real marginal cost. The presence of search and matching frictions changes the nature of real marginal cost because it affects both the way the labor input is used to produce output and the way its price is determined.

First, recall that the relative price of intermediate goods x_t coincides with the real marginal cost faced by retailers. Then, in order to shed some light on the determination of the real marginal cost, note that final goods output is produced according to equation (32):

$$y_t = \overline{n}_t h_t,$$

which says that output can be changed either by changing the number of hours that each employed worker supplies, h_t, or by changing the number of employed workers producing in t, $\overline{n}_t = n_t(1-\rho_t)$. Since it takes time to hire a new worker through vacancy posting and matching, i.e., the stock of currently employed workers, n_t, is predetermined, the number of producing workers at time t can only be changed by varying the rate at which currently employed workers are separated from their jobs, ρ_t. In other words, output can be raised to meet a higher demand either by increasing the number of hours supplied, h_t, or by raising the number of employed workers producing in t through a reduction in the job destruction rate, ρ_t. This implies that the hours per worker and the endogenous job destruction rate adjust to meet demand so that in equilibrium their marginal costs are equalized.

[21] See Rotemberg and Woodford (1997) and Boivin and Giannoni (2005) for a brief discussion.

The optimal condition for the determination of hours (22) gives the cost of producing one additional unit of output by increasing hours worked:

$$x_t = \frac{g_h(h_t, a_t)}{f_h(h_t)\lambda_t},$$

which equals the marginal disutility of supplying one additional hour, normalized by the marginal utility of consumption, divided by the marginal product of labor. It is important to note that while this condition turns out to have exactly the same form as in a neoclassical labor market, there are two important differences. First, in a neoclassical labor market, the wage adjusts to equate the value of the marginal product of hours to the marginal rate of substitution. In contrast, with privately efficient bargaining and search and matching frictions, the wage plays no allocational role for hours, meaning that a firm and a worker determine hours by equating costs and benefits of additional hours for the match as a whole. The wage only represents a transfer between the worker and the firm and needs not to be equal to neither the value of the marginal product nor to the marginal rate of substitution. Second, condition (22) in the current model determines hours per worker, the intensive margin, as opposed to total hours in the neoclassical framework. Because in the model with search most of the adjustment of the labor input occurs at the extensive margin, this has the important implication that, other things being equal, the marginal cost will change significantly less than in the baseline model where all variation in total hours occurs at the intensive margin. Thus, the ability of the model to reproduce the important feature of the data that the labor input responds to monetary policy shocks mostly at the extensive margin, also makes the model better able to account for the sluggish response of inflation to monetary policy shocks.

The analysis of what determines the marginal cost of producing output by changing the job destruction rate is specific to this model and new in the literature. It can be illustrated by interpreting the condition that determines endogenous job destruction, equation (24), which can be rearranged as

$$x_t = \frac{1}{h_t}\left(b + \frac{g(h_t, \underline{a}_t)}{\lambda_t}\right) - \frac{1-\eta s_t}{1-\eta}\frac{\kappa}{\lambda_t q_t h_t}$$

The cost of producing one additional unit of output by increasing employment through a reduction in the job destruction rate has two components. The first is the flow cost (per hour) of keeping one additional worker employed: the foregone unemployment benefit plus the disutility from supplying hours of work. Note that the additional worker is one who survives at the margin. That is, the realization of his preference shock corresponds to the threshold value $\underline{a}_t$. The second term represents the (per hour) expected future joint payoffs from continuing the relationship in the following periods. Because firms and workers have long term employment relationships, the marginal cost is given by the flow cost net of the expected future payoffs associated with preserving an additional match through a reduction in separations. It is less straightforward to use equation (24), as opposed to equation (22), to compare the behavior of marginal cost with the standard NK model. To gain some intuition, note that in a neoclassical labor market the marginal cost would simply be equal to the marginal rate of substitution. Similarly, here the first component of marginal cost is positively related to the labor supply disutility. However, since expected future payoffs from continuing the relationship are procyclical, the response of x_t to a monetary policy shock is dampened by the presence of the second component.[22] Note finally that with Nash bargaining the wage does not play any allocational role for the decision to continue or separate a match, or the decision of how many hours to work.[23]

5 Bringing the model to the data

In this section I describe the econometric methodology that I use to evaluate the model developed in Section 3. The model parameters can be divided in three groups. The first group is composed by the parameters that characterize the Taylor rule and is given by $\left\{\rho_m, \gamma_\pi, \gamma_y\right\}$. The second group is given by the structural parameters that affect the dynamics of both the search model and the baseline NK model. This group is given by $\{\beta, \phi, \kappa_h, e, \varepsilon, \psi, \omega\}$. The third group includes the structural parameters that describe the labor market in the search model. This group does not affect the dynamics of the standard NK model and is composed by $\{\alpha, \sigma, \eta, \mu_a, \sigma_a, \rho, q, n\}$.[24]

First, I set the Taylor rule parameters as follows: the interest rate smoothing parameter ρ_m is set to be equal to 0.85, and the parameters γ_π and γ_y to 1.5 and 0.5, respectively. These values are roughly consistent with the estimates presented in Clarida, Galí and Gertler (2000).

Second, I calibrate the parameters of the second group, with the exception of the habit persistence parameter e. Specifically, I set the quarterly discount factor β to 0.99, which implies a quarterly real rate of interest of approximately 1 percent. In order to calibrate the parameter ϕ, note first that $1/\phi$ is the intertemporal elasticity of substitution of leisure. The value of this elasticity has been a substantial source of controversy in the literature. Most microeconomic studies estimates this elasticity to be small, close to 0 and not higher than 0.5.[25] Scholars of the business cycle, however, tend to work with elasticities that are much higher than microeconomic estimates, typically unity and above. In such a way they can approximate the absence of the extensive margin variation of the labor input. Since the model that I develop in this paper can account for both margins, I accordingly set ϕ equal to 10, which implies an elasticity of intertemporal substitution of 0.1. I set the probability that a firm does not change its price within a given period, ψ, equal to 0.85, implying that the average period between price adjustments is around 6.5 quarters. The

[22] Since the firm and the worker share the joint surplus according to the parameter η, $S_t(\underline{a}_t) = 0$ is equivalent to $J_t(\underline{a}_t) = 0$. Substituting and rearranging, then, one could write marginal costs as $x_t = w_t\left(\underline{a}_t\right) - \frac{\kappa}{\lambda_t q_t h_t}$. In this case, the marginal cost is the hourly wage paid to the additional worker (evaluated at the threshold value of the preference shock) minus the savings per hour in terms of expected costs of hiring a worker, given by $\frac{\kappa}{\lambda_t q_t h_t}$.

[23] Expressing the job destruction condition in terms of the wage, as in previous footnote, may then be misleading.

[24] To be precise, ρ, q and n are targeted steady state values. Accordingly, the parameters σ_m, κ and b are calculated from steady state relationships.

[25] For a survey of the literature see Card (1994).

fraction ω of backward-looking retailers is set to 0.5. Both values are consistent with the estimates in Galí and Gertler (1999).[26] I assume that the markup of prices on marginal costs is on average 10 percent. This amounts to setting ε equal to 11. Finally, I normalize the time spent working in the steady state, h, to 1 and set κ_h accordingly.

Third, I estimate a set of the structural parameters that characterize the labor market in the search model. Moreover, since the habit persistence parameter is important to explain the dynamics of the labor market, I include it in the group of parameters to be estimated. The following two sections describe the estimation procedure and results.

5.1 Minimum distance estimation

I follow the estimation strategy adopted in Rotemberg and Woodford (1997), Christiano, Eichenbaum and Evans (2005) and Boivin and Giannoni (2005) in using a limited information minimum distance estimator. Specifically, the parameters are chosen so that the impulse responses to the monetary policy shock of a set of endogenous variables in the model match as closely as possible the responses estimated from the VAR.

More formally, denote with Ψ the vector of labor-market parameters to be estimated and with $g_M(\Psi)$ the vector-valued function containing the model-based impulse response functions. Then, denote with Φ the vector of the estimated VAR coefficients and with $g_V(\Phi)$ the vector-valued function containing the VAR-based impulse response functions. The minimum distance estimator, $\widehat{\Psi}$, can be obtained by minimizing the objective function

$$L(\Psi) = [g_M(\Psi) - g_V(\Phi)]' \Lambda [g_M(\Psi) - g_V(\Phi)],$$

with respect to Ψ and subject to the theoretical constraints on the values of the parameters. In the objective function, Λ denotes a diagonal weighting matrix with the inverse of each impulse response's variance along the diagonal. The choice of this weighting matrix effectively takes into account that some of the point estimates of the impulse responses are less accurate than others. Finally, I consider in the estimation the impulse responses of the variables y_t, π_t, r_t^n, $\overline{n}_t$, h_t, jc_t and jd_t over the first twenty periods after the monetary policy shock.

As Dridi and Renault (2001) and Boivin and Giannoni (2005) point out, although this estimation strategy is similar in the spirit to a calibration exercise, it produces consistent estimates of the structural parameters on which it is possible to perform statistical inference. Moreover, given that the main purpose of this study is to explain the response of the economy to a monetary policy

[26] It is important to point out that it is not necessary to rely on such high values of the parameters φ and ω to explain inflation dynamics in the data. In particular, I could allow for some form of real rigidity and, everything else equal, significantly reduce the value of both parameters. However, for clarity of presentation and analogously to Galí and Gertler (1999), I do not include in the model this additional feature. Moreover, as I discuss below, the important result is that for given values of φ and ω, whichever values I assume, the response of inflation is smaller than in the baseline NK model.

shock, the estimation based on the impulse responses permits me to focus on the moments of the data that the model seeks to explain.

Among the labor market parameters and steady state values, three of them can be easily calibrated from the data. In particular, the empirical literature provides us with several measures of the U.S. worker separation rate. Davis, Haltiwanger and Schuh (1996) compute a quarterly worker separation rate of about 8 percent, while Hall (1995) reports this rate to be between 8 and 10 percent. Accordingly, I set the overall separation rate ρ to 0.08. In order to calibrate α, I follow den Haan, Ramey and Watson (2000). First, as previously discussed, they assume that only exogenous separations are reposted. Then, based on evidence reported by Davis, Haltiwanger and Schuh, they calculate that the rate at which separations are reposted by firms is equal to 0.68. This implies that $\alpha = 0.68$ and $\rho^x = 0.054$. Then, I set the steady state probability that a firm fills a vacancy, q, to be equal to 0.7, as in Cooley and Quadrini (1999) and den Haan, Ramey and Watson (2000). This value implies that the average time until a vacancy is filled is 1.4 quarters. The vector of parameters and steady state values to be estimated, then, is given by $\Psi = [\sigma, \eta, n, e, \mu_a, \sigma_a]$. Below I discuss why I choose to estimate the steady state employment rate n.

5.2 Estimation results

I perform the estimation in two stages. In the first stage I estimate all six parameters $\sigma, \eta, n, e, \mu_a$ and σ_a. The relative bargaining power η and the lognormal parameter μ_a are very imprecisely estimated. This suggests that they may not have a large effect on the dynamics of the model. I then set the lognormal parameter μ_a to 0 and the relative bargaining power parameter η to 0.5 and estimate σ, n, e and σ_a. The estimates of these parameters are the same as those in the first stage, only the standard errors are smaller, confirming that η and μ_a have a negligible impact on the dynamic behavior of the model. The final estimates of σ, n, e and σ_a are all statistically significant and are reported in Table 1, along with the corresponding standard errors.

Table 1: Estimation results

σ	η	n	e	μ_a	σ_a
0.545	0.5	0.747	0.55	0	0.382
(0.2839)	$(-)$	(0.1022)	(0.0562)	$(-)$	(0.0868)

Note: Standard errors are in parenthesis. $(-)$ denotes that the standard error is not available because the parameter is calibrated in the final estimates.

The elasticity of new matches with respect to the number of searching workers, σ, is estimated to be 0.55. This value is higher but not too far from the estimate of 0.4 obtained by Blanchard and

Diamond (1989) and it is consistent with the evidence summarized by Petrongolo and Pissarides (2001) in their survey on the matching function.

The estimate of the habit persistence parameter, e, is 0.55. This is close to the estimate of 0.63 reported in Christiano, Eichenbaum and Evans (2005).

The reason why I choose to estimate the steady state employment ratio n is that on one hand it may have considerable effects on the dynamics of the labor market, on the other there is no unambiguous way to calibrate it from the data. More precisely, as an example, Andolfatto (1996) sets the employment rate n to 0.54, while den Haan, Ramey and Watson (2000) set it to 0.89. These values, which are obviously lower than in the data, can be justified by interpreting the unmatched workers in the model as being both unemployed and partly out of the labor force. This interpretation is consistent with the abstraction in the model from labor force participation decisions. The estimate of n that I obtain is 0.75. This estimate lies between the values just cited.

The estimate of the parameter σ_a is 0.38. Along with the parameter μ_a, this value determines the steady state value of the threshold of the idiosyncratic shock, $\underline{a}$, and the elasticity of output to changes in the threshold value, $\eta_{F,\underline{a}}$, from the steady state relationships. The implied values for $\underline{a}$ and $\eta_{F,\underline{a}}$ are, respectively, 2.2 and 0.17.

Finally, given the above estimates, the steady state probability that a worker finds a job, s, is calculated from the steady state relationships to be 0.2, implying that the average time until a worker finds a job is 5 quarters. Recall that in the model we have interpreted the pool of searching workers as both unemployed and partly out of the labor force. The parameters κ and b are also derived from the steady state calculation and imply, respectively, a vacancy cost to output ratio of about 0.05 and a replacement rate of about 0.1.

6 Findings

In order to assess the contribution of labor market frictions to shaping the economy dynamics, I start by comparing the predictions of the model developed in this paper - from now on the search model - with those of a baseline NK model with a neoclassical labor market. The latter is obtained from the search model simply by assuming away search and matching frictions or, equivalently, by shutting down endogenous job creation and destruction. All other features - frictions, parameters and information lags - are kept unchanged. Therefore, any difference in the dynamics of the economy and, in particular, in the predicted joint behavior of output and inflation in the two models is associated with the presence of labor market frictions.

Figure 3 shows the response of output, inflation, marginal costs and hours to a one percent increase in the nominal interest rate in both models.[27] All variables have a similar qualitative response in the search and the baseline NK model. The tightening in monetary policy reduces output of final goods, hours worked, marginal cost and inflation.

[27] The figure shows the annualized inflation and nominal interest rates.

From a quantitative point of view, however, the search and the baseline model behave quite differently. First, the search model generates a much lower volatility of inflation relatively to the volatility of output. While in the baseline model a peak decrease in output of about 0.23 percent is associated with a peak fall in inflation of around 0.63 percent, in the search model output falls by about 0.33 percent and inflation by only 0.27 percent. More precisely, the ratio of the inflation standard deviation to the output standard deviation is 0.74 in the baseline model and 0.20 in the search model. The relative volatility in the search model is very close to the value of 0.17 found in the data (see Table 2 below). Second, the search model generates a significantly more persistent response of output. In the baseline model output goes back to its steady state value after 9 quarters, while in the search model it takes around 20 quarters. In order to quantify more precisely the increase in output persistence, I calculate that in the search model 68 percent of the cumulative response of output occurs after the fourth quarter and 18 percent after the eighth quarter. In the baseline model, instead, only 46 percent of the cumulative decrease occurs after the fourth quarter and almost 0 percent after the eight quarter. I also calculate the ratio of the amount of time when the response of output is negative to the number of periods in a typical contract. Following Chari, Kehoe and McGrattan (2000) and Christiano, Eichenbaum and Evans (2005), I call this number the contract multiplier. This statistic is equal to 1.05 in the baseline model and to 2.7 in the search model. Third, as an additional way to assess the contribution of labor market frictions in accounting for output and inflation dynamics, I calculate that the search model would generate responses of output and inflation similar to those in the baseline model if the average time a firm keeps prices fixed is reduced from around 6.6 quarters to a much lower value of 3.3 quarters, everything else equal.

The lower volatility of inflation relative to output and the larger persistence of output are caused by the substantially lower elasticity of marginal cost with respect to output. The figure shows that a given fall in output is associated with a much lower decrease in the level of marginal cost than in the baseline model. In turn, smaller variations in marginal cost induce firms setting their prices to make smaller adjustments in prices. This increases the sluggishness of the aggregate price level to changes in aggregate demand and reduces the volatility of inflation. Finally, the lower sensitivity of the price level to variations in aggregate demand raises the persistence of the response of aggregate demand and output to a monetary policy shock.

The elasticity of marginal cost with respect to output is lower in the search model because the labor input can vary at both the intensive and the extensive margin. To see this, combine the loglinear version of equations (22) and (32) to obtain $\widehat{x}_t = \phi(\widehat{y}_t - \widehat{\overline{n}}_t) - \widehat{\lambda}_t$ with $\widehat{\overline{n}}_t = \widehat{n}_t + \eta_{F,\underline{a}}\widehat{\underline{a}}_t$, where variables with a "hat" are log deviations from the steady state. For a given change in output, the change in marginal cost is smaller the larger is the share of the fluctuation in total hours that takes the form of fluctuations in employment, through changes in the job destruction rate. Figure 4 plots the responses of total hours, employment (after job destruction has taken place) and hours per worker in the search model. The percent change in total hours is the sum of percent changes in employment and hours per worker. The figure shows that the decrease in the number of people working is substantially larger and more persistent than the fall in hours per worker. Initially, the fall in the demand for intermediate goods reduces its relative price and reduces hours per worker. At the same time, the lower profitability of firms induces fewer firms to post vacancies and more firms to separate from their workers. As the number of intermediate goods firms producing gradually decreases, the demand of intermediate goods per firm gradually increases. As a consequence, the responses of output per firm and hours of work in the intermediate goods sector are reverted fairly quickly.

It must be emphasized that I have assumed a degree of intertemporal substitution in the supply of hours that is consistent with microeconomic estimates. The intertemporal elasticity of substitution of leisure is assumed to be equal to 0.1. Instead, general equilibrium models of the business cycle, among which sticky prices models, tend to assume much higher values of this elasticity, typically unit and above. By doing so, they can approximate some implications of the model with both margins of adjustment. In particular, the baseline NK model that I consider can approximately replicate the joint behavior of output and inflation in the search model if the elasticity is increased from a value of 0.1 to values between 1.5 and 2, everything else unchanged. Of course, such a model cannot explain what drives fluctuations in employment as opposed to hours per worker, why there is unemployment in equilibrium or, more generally, it cannot explain the behavior of the labor market over the business cycle.

Figure 5 presents the dynamics of the labor market in the search model after a monetary policy shock. The response of employment is explained by the dynamics of job creation and job destruction. Recall, from equation (27), that employment growth is given by $\frac{n_{t+1}-n_t}{n_t} = jc_t - jd_t$. As can be seen from the figure, a contractionary monetary policy shock decreases job creation and raises job destruction. The increase in job destruction is slightly greater and significantly more persistent than the decrease in job creation. Thus, most of the decrease in employment is due to the response of job destruction.

The responses of job creation and destruction, in turn, can be explained as follows. A persistent increase in the nominal interest rate causes a decrease in current and expected future aggregate demand. The fall in aggregate demand, in turn, decreases the demand for intermediate goods and the profits of firms producing them. This diminishes the value of the idiosyncratic shock above which the firm and the worker decide to separate and raises the separation rate. The decrease in profits also reduces the value of opening a vacancy and induces firms to post less vacancies, yielding a fall in the job creation rate. Because of the timing assumption, the monetary policy shock only affects job creation and destruction with a two-period delay.

The decrease in the number of posted vacancies and the increase in the number of searching workers cause the labor market tightness to decrease. Thus, the probability of filling a vacancy increases while the probability of finding a job drops. The higher probability of hiring a worker increases the attractiveness of hiring activities and the expected future value of a match. Therefore, job creation starts to increase and job destruction to fall.

Figure 6 plots the model impulse responses of output, inflation and the nominal interest rate to the monetary policy shock against the estimated impulse responses in the U.S. economy. Figure 7 plots the responses of employment, hours per worker, job creation and job destruction. The solid and dashed lines denote, respectively, the estimated impulse responses and the two standard deviations confidence intervals, while the lines with circles denote the simulated responses in the model. The model does a good job in accounting for the dynamic response of the U.S. economy to a monetary policy shock.

The first dimension in which the model can reproduce the data is the joint dynamic behavior of output and inflation. Basically, the simulated responses of output and inflation are everywhere within the respective confidence intervals. However, while the model generates significantly more persistence in output than the baseline NK model, Figure 6 suggests that output is not yet as persistent as in the data. Second, the model is able to reproduce the quantitative behavior of the variation of the labor input at both margins of adjustment. It generates a small, transitory fall in hours per worker together with a larger, more persistent fall in employment. Likewise the response of output, however, the response of employment is less persistent than in the data. Third, the model explains the joint behavior of job creation and job destruction. In particular, it can account for the larger response of job destruction than job creation and for the observed upturn in job creation. This upturn occurs because the larger pool of unemployed workers looking for a job stimulates firms to post new vacancies. The model can also account for the higher degree of persistence in job destruction with respect to job creation that is observed in the data. Note that the simulated impulse responses of all four labor market variables are everywhere within the respective confidence intervals.

Table 2: Relative standard deviations

	σ_{π}/σ_y	$\sigma_{\overline{n}}/\sigma_y$	σ_h/σ_y	σ_{jc}/σ_y	σ_{jd}/σ_y
U.S. Data	0.17	0.89	0.20	5.10	7.68
Model	0.20	0.88	0.16	4.62	5.31

Table 2 reports the ratio of the standard deviation of π_t, $\overline{n}_t$, h_t, jc_t and jd_t to that of output y_t, conditional on the monetary policy shock, both in the model and in the data. The model is quite successful at replicating the relative standard deviations of the endogenous variables on which the analyses is focused, even though the standard deviation of most labor market variables is somewhat below the respective values in the data.

Finally, I study the role of habit formation in consumption preferences in explaining the response of the economy to a monetary policy shock. A number of authors have argued that habit persistence is important to understand the transmission mechanism of monetary policy shocks.[28] In particular,

[28] See for example McCallum and Nelson (1999), Fuhrer (2000) and Christiano, Eichenbaum and Evans (2005).

it helps to account for the hump-shaped persistent decrease in output together with the rise in the real interest rate after a contractionary monetary policy shock. This is also true in the model developed here, for two reasons: output is demand-determined, as in the NK model without search frictions, and I only consider demand shocks.[29] In addition, in the search model habit formation also helps to account for the observation that the labor input varies mostly at the extensive margin. The intuition is the following. Without habit persistence, because the model cannot generate a hump-shaped response in consumption, the largest change in aggregate demand and output occurs in the first period of output variation. Then, since employment only moves gradually, hours per worker have to vary significantly to accommodate changes in demand-determined output. To illustrate this effect, I set the habit persistence parameter e to 0 in the search model, keeping the value of all other parameters unchanged. I obtain that the standard deviation of employment relative to output, $\sigma_{\overline{n}}/\sigma_y$, decreases from 0.88 to 0.76, while the standard deviation of hours per worker relative to output, σ_h/σ_y, increases from 0.16 to 0.33. That is, in the absence of habit persistence a relatively larger share of the change in total hours is counterfactually accounted for by changes at the intensive rather than the extensive margin.

I then compare the search model and the baseline NK model assuming no habit persistence in both models. First, the search model continues to generate a lower volatility of inflation relative to output: the ratio of inflation standard deviation to output is 0.42 in the NK model and 0.14 in the search model. Second, the response of output is still more persistent in the search model than in the NK model: in the search model 47 percent of the cumulative response of output occurs after the fourth quarter and 10 percent after the eight quarter, while in the NK model only 19 percent of the cumulative decrease occurs after the fourth quarter and none after the eight quarter. That is, while search frictions cannot by themselves generate the hump-shaped response of output that is observed in the data after a monetary policy shock, they contribute to its persistence.

7 Search frictions versus nominal wage rigidities

The recent papers by Smets and Wouters (2003) and Christiano, Eichenbaum and Evans (2005) show that when staggered nominal wage setting is included, along with a number of other features also present in this paper, a more conventional NK model without search can also match the responses of output and inflation to a monetary policy shock. These papers, however, have employment adjusting along the intensive margin. That is, wage stickiness affects fluctuations in hours worked as opposed to total employment. As a consequence, these frameworks are subject to Barro's (1977) argument that wages may not be allocational in this kind of environment, given that firms and workers have an on-going relationship. If wages are not allocational, of course, then wage rigidity does not influence the dynamics of the model. For this reason, relying on staggered nominal wage setting with employment varying at the intensive margin to explain inflation persistence and

[29] Output is solely determined by consumption demand as I abstract from capital and government spending.

output dynamics is not completely satisfactory.

However, it may still be interesting to document how the model with nominal wage rigidities compares to the model with search frictions in terms of the response of output and inflation to monetary policy shocks. To do this, I augment the baseline NK model with two variants of the Calvo-type nominal wage stickiness adopted in the literature. One in which each household is the monopolistic supplier of a differentiated type of labor input, as in Erceg, Henderson and Levin (EHL, 2000) and one in which households supply an homogeneous labor input that is transformed by monopolistically competitive labor unions into a differentiated labor input as in Schmitt-Grohe and Uribe (SGU, 2006a).[30]

In such models the log-linear version of the wage Phillips curve is given by

$$\widehat{w}_t = \frac{1}{1+\beta}\left(\widehat{w}_{t-1} + \widehat{\pi}_{t-1} - \widehat{\pi}_t\right) + \frac{\gamma}{1+\beta}\left(\phi\widehat{h}_t - \widehat{\lambda}_t - \widehat{w}_t\right) + \frac{\beta}{1+\beta}E_t\left(\widehat{w}_{t+1} + \widehat{\pi}_{t+1} - \widehat{\pi}_t\right), \qquad (34)$$

where the parameter γ differs across the two variants as follows:

$$\gamma^{EHL} = \frac{(1-\beta\psi^w)(1-\psi^w)}{\psi^w}\frac{1}{1+\varepsilon^w\phi} \quad \text{and} \quad \gamma^{SGU} = \frac{(1-\beta\psi^w)(1-\psi^w)}{\psi^w},$$

with ε^w denoting the elasticity of substitution across differentiated labor inputs and $1-\psi^w$ the probability with which each household (or labor union) can reset its wage in any given period.[31]

I then compare the responses of output and inflation to the monetary policy shock in the search model with the responses of the same variables in the two variants of the sticky wage model for different degrees of wage stickiness.[32] Figure 8 plots the results. Not surprisingly, the responses of output and inflation depend on the degree of wage stickiness, as well as on the type of Calvo wage stickiness. In particular, the figure shows that inflation in the search model is less volatile than in the SGU model and more volatile than in the EHL model at all degree of wage stickiness. To shed additional light on the comparison between the search and the sticky wage model, I then compute the ratio of the inflation standard deviation to output in the EHL and SGU models. Recall that this ratio is 0.17 in the data and 0.20 in the search model.

With the current parameterization, the SGU model generates too much volatility in inflation. At the lower degree of wage rigidity, with $\psi^w = 0.5$, the relative volatility of inflation is not much different from the relative volatility generated by the baseline NK model with no labor market frictions (0.67 versus 0.74). As the degree of wage rigidity increases, inflation volatility in the SGU model decreases, but even with wages fixed on average for 6.5 quarters the model generates twice the volatility that is observed in the data (0.38 in the model versus 0.17 in the data). The EHL model, instead, gets close to the data when wages are fixed for 2 to 3 quarters on average. The EHL

[30] See Schmitt-Grohe and Uribe (2006b) for a comparison of the two types of Calvo nominal wage stickiness.

[31] The wage Phillips curve, together with an equation equating the value of the marginal product of hours, x_t, to the real wage, w_t, determines optimal hours.

[32] The parameter ε^w is set so that the markup of wages on the marginal rate of substitution is on average 10 percent and I consider values of 0.5, 0.66 and 0.85 for the parameter φ^w, implying that wages are fixed on average for 2, 3 and 6.5 quarters, respectively.

model with $\psi^w = 0.5$ is the closer one to the search model: the responses of output and inflation are almost the same with inflation slightly less persistent and output slightly more persistent in the search model. This suggests that the EHL model could fit the data at a higher, but reasonable, degree of wage stickiness allowing at the same time for a lower degree of price stickiness. Finally, note that while the EHL model is not very sensitive to changes in the value of the intertemporal elasticity of substitution of leisure, $1/\phi$, the SGU model predicts lower ratios of inflation to output standard deviation when this elasticity is higher.

Table 3: Inflation to output standard deviation

	EHL	SGU
$\psi^w = 0.5$	0.20	0.67
$\psi^w = 0.66$	0.13	0.59
$\psi^w = 0.85$	0.05	0.38

A number of recent papers, beginning with Shimer (2005) and Hall (2005), consider the role of real wage rigidity in explaining labor market dynamics within a baseline Mortensen and Pissarides model.[33] By enhancing the cyclicality of firms' profits and incentives to hire, wage stickiness helps to account for the volatility of labor market activity that is observed in the data. Moreover, as emphasized by Hall (2005), because wage rigidity affects employment at the extensive margin, in these frameworks the Barro's critique does not apply. These analyses differ from the one conducted in this paper in several aspects: they study non-monetary models, only consider the extensive margin and, finally, evaluate the model against unconditional moments taking technology shocks as the exogenous force driving fluctuations. Nevertheless, they suggest that wage rigidities may improve the performance of the present model by raising the volatility of labor market activity, conditional on a monetary policy shock. In addition, they may reduce the need to rely on frictions such as habit persistence and price stickiness to account for the cyclical behavior of output and inflation. In particular, it seems promising to incorporate in the current framework the model of wage rigidity developed by Gertler and Trigari (2006), who modify a baseline Mortensen and Pissarides model to allow for multiperiod staggered wage setting. This is part of an on-going research project in Gertler, Sala and Trigari (2007).

8 Conclusions

This paper builds on the NK theory of money and inflation and the modern theory of equilibrium unemployment. Both theories have been introduced previously in the macroeconomic literature and

[33] See Hall (2005) for a survey.

extensively used for both normative and positive analysis. But the combination of these theories into a single dynamic general equilibrium model provides new insights on the linkages between monetary policy, business cycle fluctuations and the dynamics of the labor market.

When labor market search is incorporated into a standard NK model, implying that changes in the labor input can occur at both the intensive and the extensive margin, the ability of the model to explain the response of output and inflation to monetary policy shocks improves along a number of dimensions. This happens because introducing search and matching frictions modifies the nature of real marginal cost faced by firms in a way that lowers the elasticity of marginal cost with respect to output. In general, the estimated model does a good job in accounting quantitatively for the response of the U.S. economy to a monetary policy shock. Moreover, the ability of the model to account for the joint dynamics of output and inflation relies on its ability to explain the dynamics in the labor market.

References

[1] Andolfatto, David (1996). "Business Cycles and Labor Market Search," *American Economic Review* 861, 112-132.

[2] Barro, Robert (1977). "Long-Term Contracting, Sticky Prices and Monetary Policy," *Journal of Monetary Economics* 3 (3), 305-316.

[3] Bernanke, Ben, Mark Gertler and Simon Gilchrist (1999). "The Financial Accelerator in a Business Cycle Framework," *Handbook of Macroeconomics*, J. Taylor and M. Woodford, editors, North-Holland.

[4] Boivin, Jean and Marc Giannoni (2006). "Has Monetary Policy Become More Effective?," *The Review of Economics and Statistics* 88(3), 445-462.

[5] Calvo, Guillermo (1983). "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics* 12, 383-398.

[6] Card, David (1994). "Intertemporal Labor Supply: An Assessment," *Advances in Econometrics*, C. Sims, editor, Cambridge University Press.

[7] Chari, Varadarajan V., Patrick J. Kehoe and Ellen R. McGrattan (2000). "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?," *Econometrica* 68(5), 1151-1180.

[8] Christiano, Lawrence, Martin Eichenbaum and Charles Evans (2000). "Monetary Policy Shocks: What Have We Learned and to What End?," *Handbook of Macroeconomics*, J. Taylor and M. Woodford, editors, North-Holland.

[9] Christiano, Lawrence, Martin Eichenbaum and Charles Evans (2005). "Nominal Rigidities and the Dynamics Effects of a Shock to Monetary Policy," *Journal of Political Economy* 113(1), 1-45.

[10] Clarida, Richard, Jordi Galí and Mark Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics* 115(1), 147-180.

[11] Cooley, Thomas and Vincenzo Quadrini (1999). "A Neoclassical Model of the Phillips Curve," *Journal of Monetary Economics* 4, 165-193.

[12] Davis Steven, John Haltiwanger and Scott Schuh (1996). Job Creation and Destruction, The MIT Press.

[13] Den Haan, Wouter , Gary Ramey and Joel Watson (2000). "Job Destruction and Propagation of Shocks," *American Economic Review* 90, 482-498.

[14] Dridi, Ramdan and Eric Renault (2001). "Semi-Parametric Indirect Inference," manuscript.

[15] Erceg, Christopher, Dale Henderson and Andrew Levin (2000). "Optimal Monetary Policy With Staggered Wage and Price Contracts," *Journal of Monetary Economics* 46, 281-313.

[16] Fuhrer, Jeffrey (2000). "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *American Economic Review* 90, 367-390.

[17] Galí, Jordi (2003). "New Perspectives on Monetary Policy, Inflation and the Business Cycle," *Advances in Economic Theory*, M. Dewatripont, L. Hansen and S. Turnovsky, editors, Cambridge University Press.

[18] Galí, Jordi (1999). "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review* 89, 249-271.

[19] Galí, Jordi and Mark Gertler (1999). "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics* 44(2), 195-222.

[20] Galí, Jordi, Mark Gertler and David Lopez-Salido (2001). "European Inflation Dynamics," *European Economic Review* 45(7), 1237-1270.

[21] Gertler, Mark and Antonella Trigari (2006). "Unemployment Fluctuations with Staggered Nash Wage Bargaining," NBER Working Paper No. 12498.

[22] Gertler, Mark, Luca Sala and Antonella Trigari (2007). "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining," manuscript.

[23] Hairault, Jean-Olivier (2002). "Labor market Search and International Business Cycles," *Review of Economic Dynamics* 5(3), 535-558.

[24] Hall, Robert (1995). "Lost Jobs," *Brookings Papers on Economic Activity* 1, 221-273.

[25] Hall, Robert (2005). "Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years," *NBER Macroeconomics Annual*, M. Gertler and K. Rogoff editors, MIT Press.

[26] Hosios, Arthur (1990). "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies* 57(2), 279-298.

[27] Ireland, Peter (1997). "A Small, Structural, Quarterly Model for Monetary Policy Evaluation," *Carnegie Rochester Conference Series on Public Policy*.

[28] Krause, Michael and Thomas Lubik (2007). "The (Ir)relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions," *Journal of Monetary Economics*, forthcoming.

[29] Levin, Andy, Alexei Onatski, John Williams and Noah Williams (2005). "Monetary Policy under Uncertainty in Micro-Founded Macroeconometric Models," *NBER Macroeconomics Annual*, M. Gertler and K. Rogoff, editors, MIT Press.

[30] McCallum, Bennett and Edward Nelson (1999). "Nominal Income Targeting in an Open-Economy Optimizing Model," *Journal of Monetary Economics* 43, 553-578.

[31] Merz, Monika (1995). "Search in the Labor Market and the Real Business Cycle," *Journal of Monetary Economics* 36, 269-300.

[32] Mortensen, Dale and Christopher Pissarides (1994). "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies* 61, 397-415.

[33] Petrongolo, Barbara and Christopher Pissarides (2001). "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* 39(2), 390-431.

[34] Rotemberg, Julio and Michael Woodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, B. Bernanke and J. Rotemberg, editors, MIT Press.

[35] Schmitt-Grohe, Stephanie and Martin Uribe (2006a). "Optimal Fiscal and Monetary Policy in a Medium Scale Macroeconomic Model," *NBER Macroeconomics Annual*, M. Gertler and K. Rogoff, editors, MIT Press.

[36] Schmitt-Grohe, Stephanie and Martin Uribe (2006b). "Comparing Two Variants of Calvo-style wage stickiness," NBER Working Paper No. 12740.

[37] Shimer, Robert (2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies: Evidence and Theory," *American Economic Review* 95(1), 25-49.

[38] Smets, Frank and Raf Wouters (2003). "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1 (5), 527-549.

[39] Trigari, Antonella (2004). "Labor Market Search, Wage Bargaining and Inflation Dynamics," IGIER Working Paper No. 268.

[40] Walsh, Carl (2005). "Labor Market Search, Sticky Prices and Interest Rate Rules," *Review of Economic Dynamics* 8, 829-849.

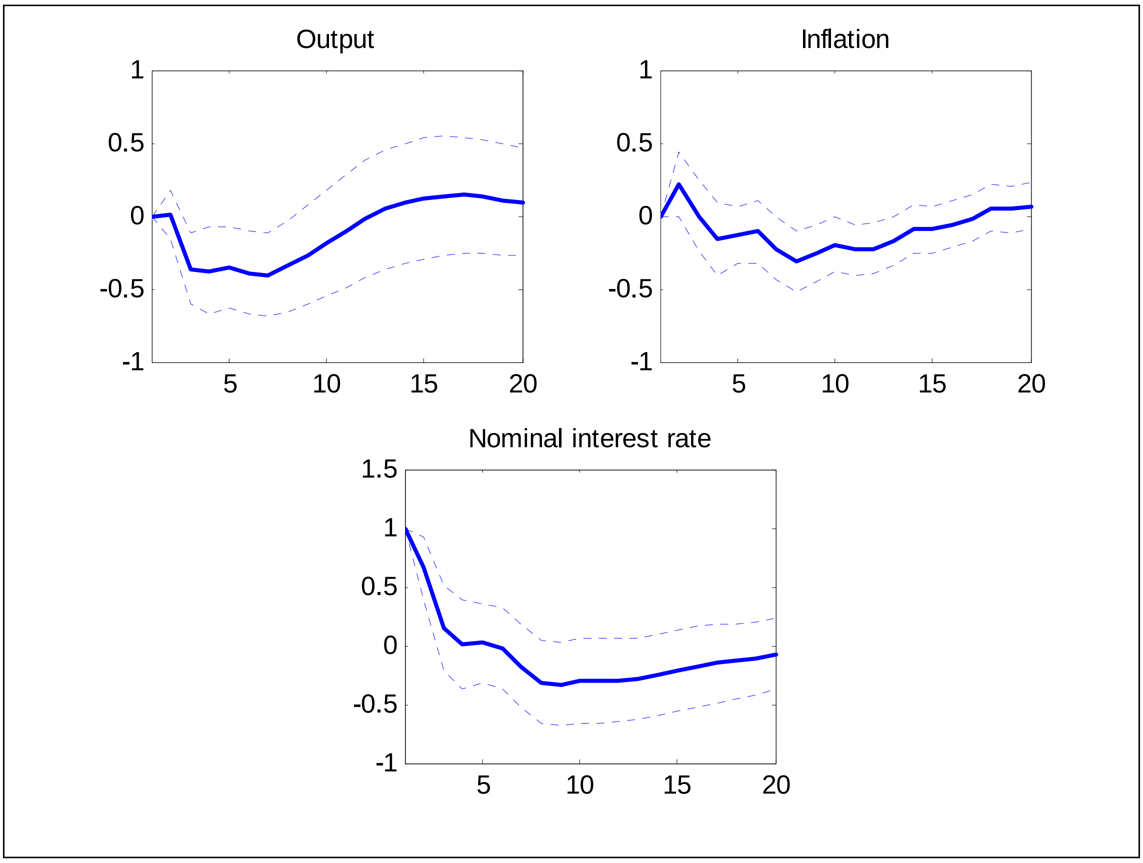

Figure 1: Estimated impulse responses to a monetary shock

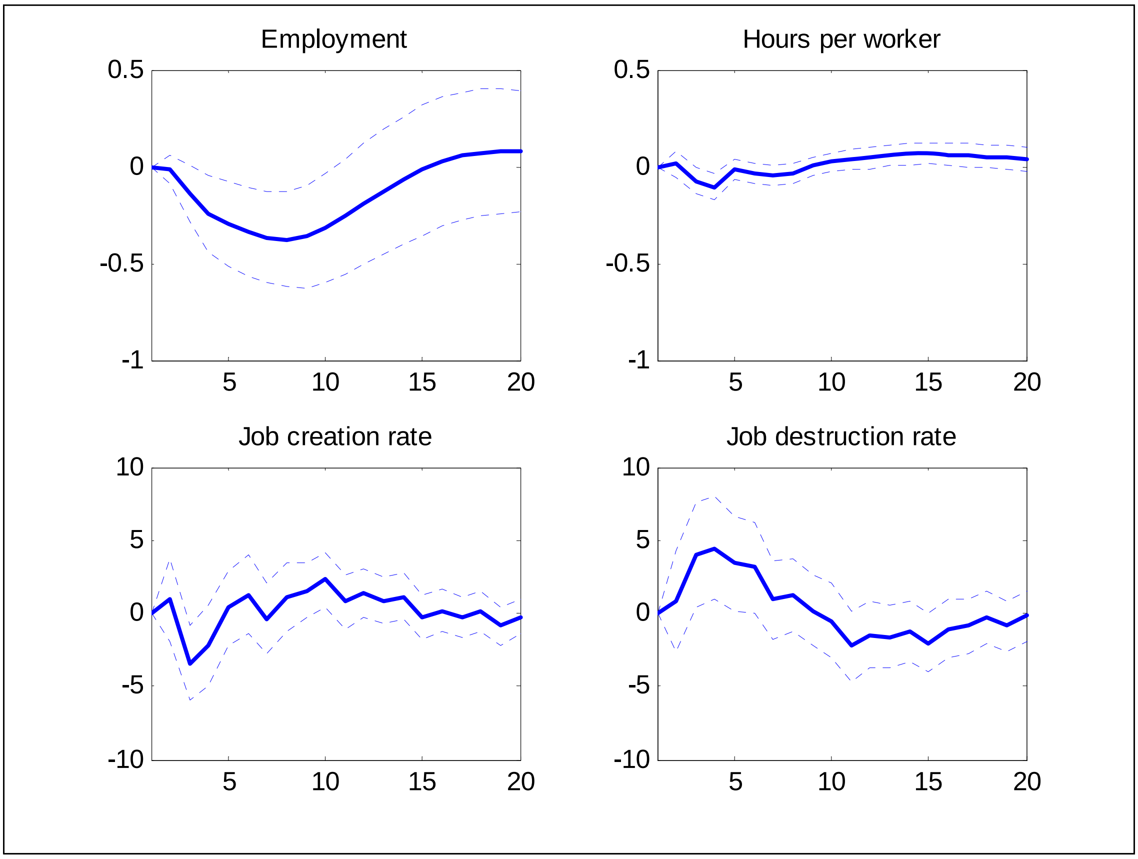

Figure 2: Estimated impulse responses to a monetary shock

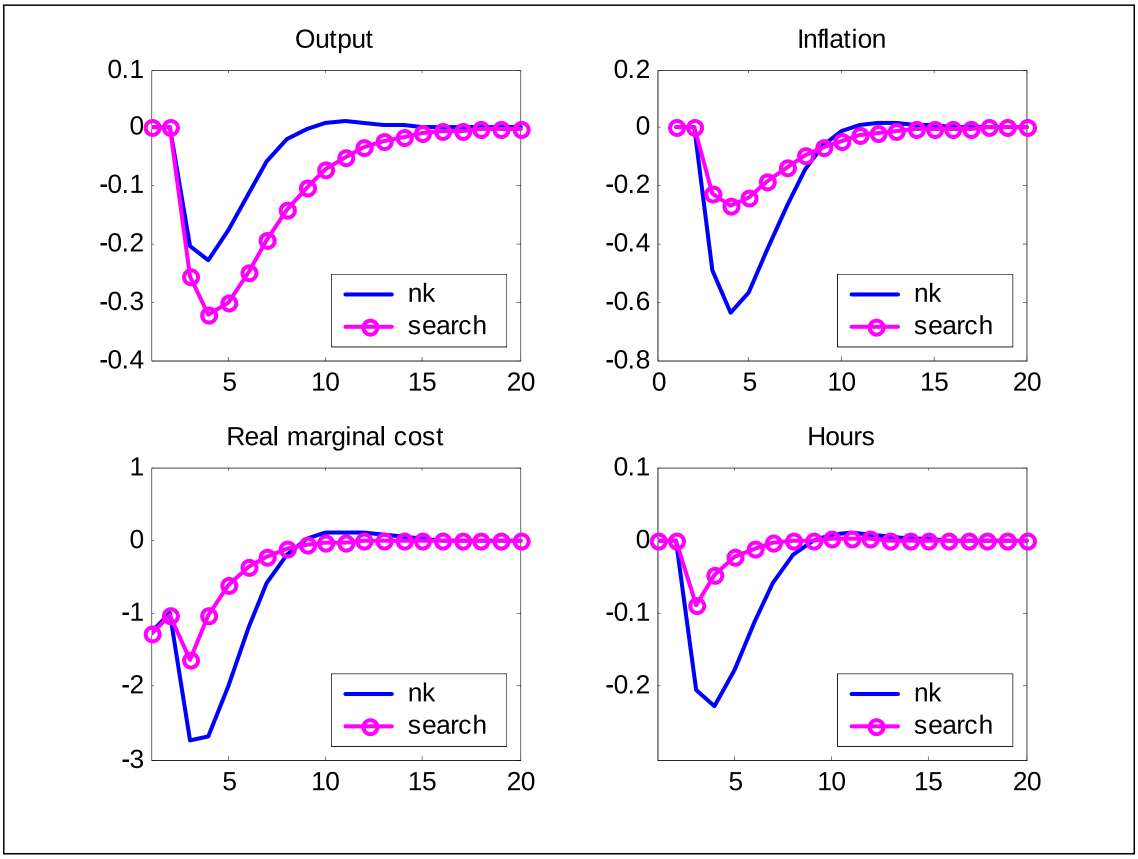

Figure 3: Search versus new keynesian model

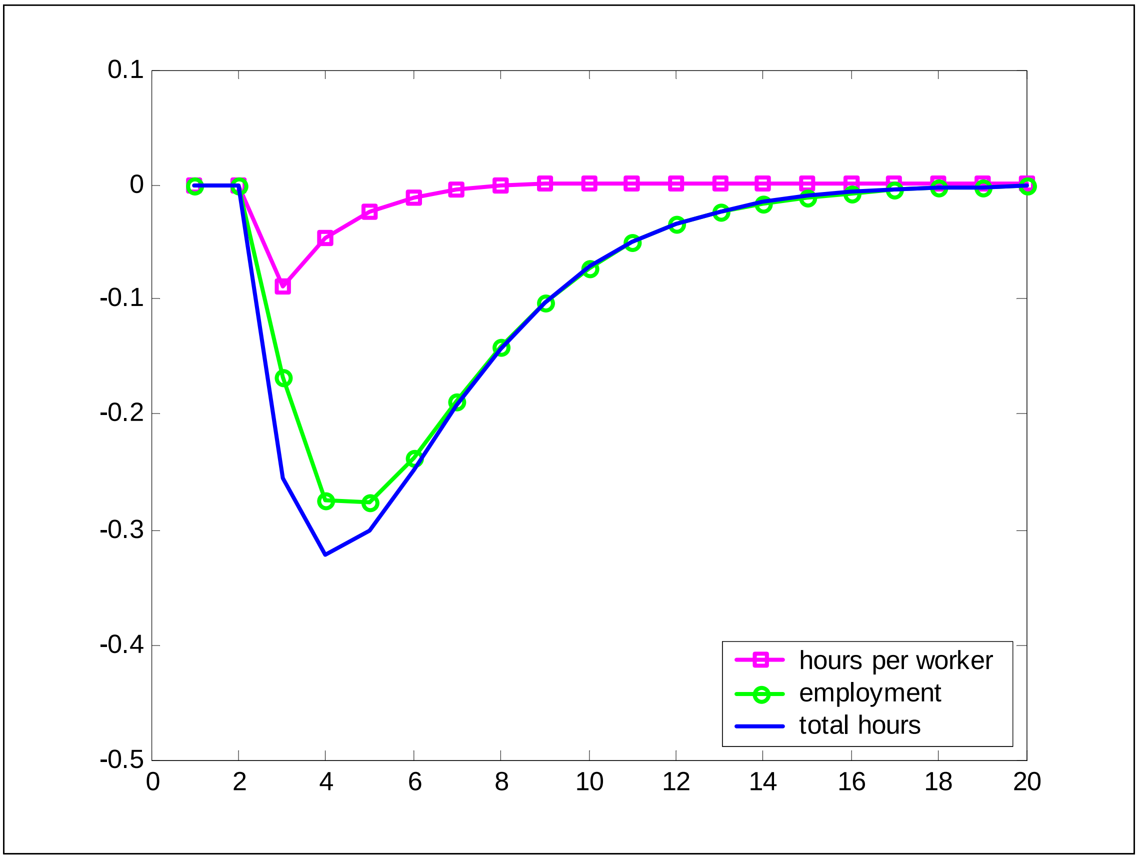

Figure 4: Extensive and intensive margin

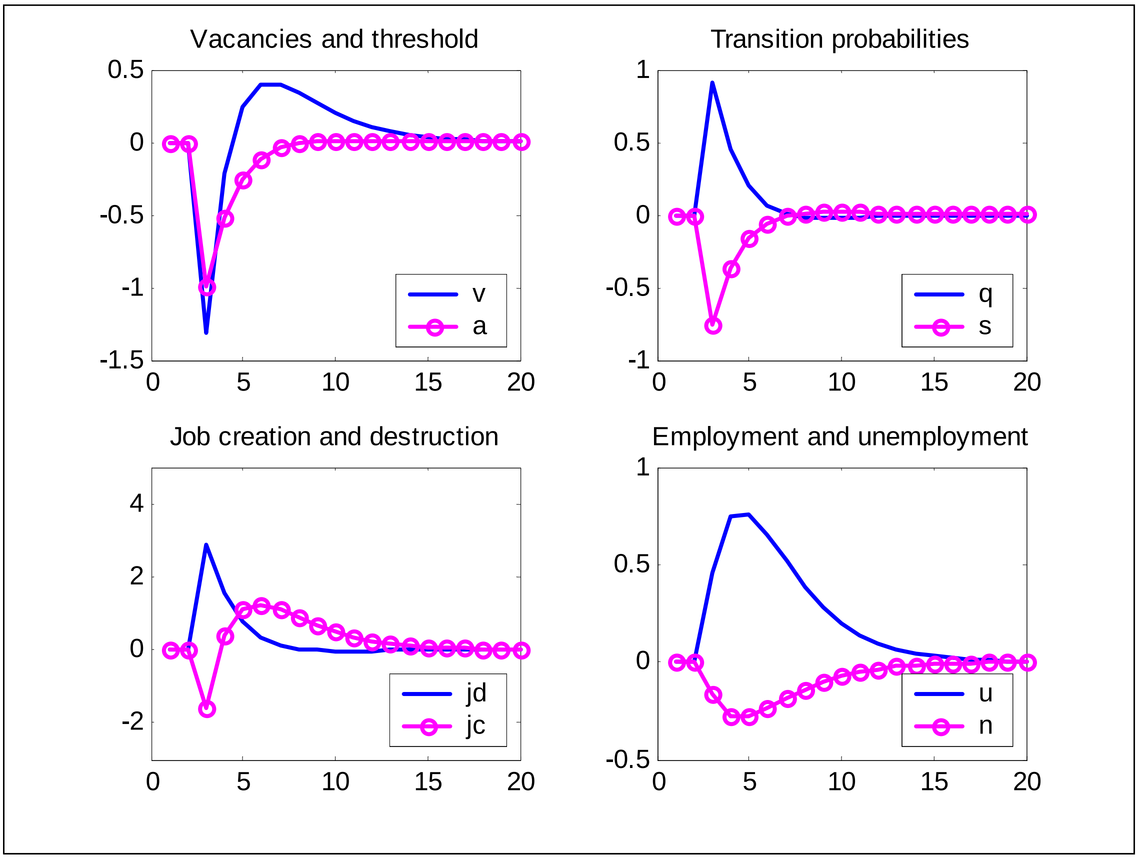

Figure 5: Labor-market dynamics

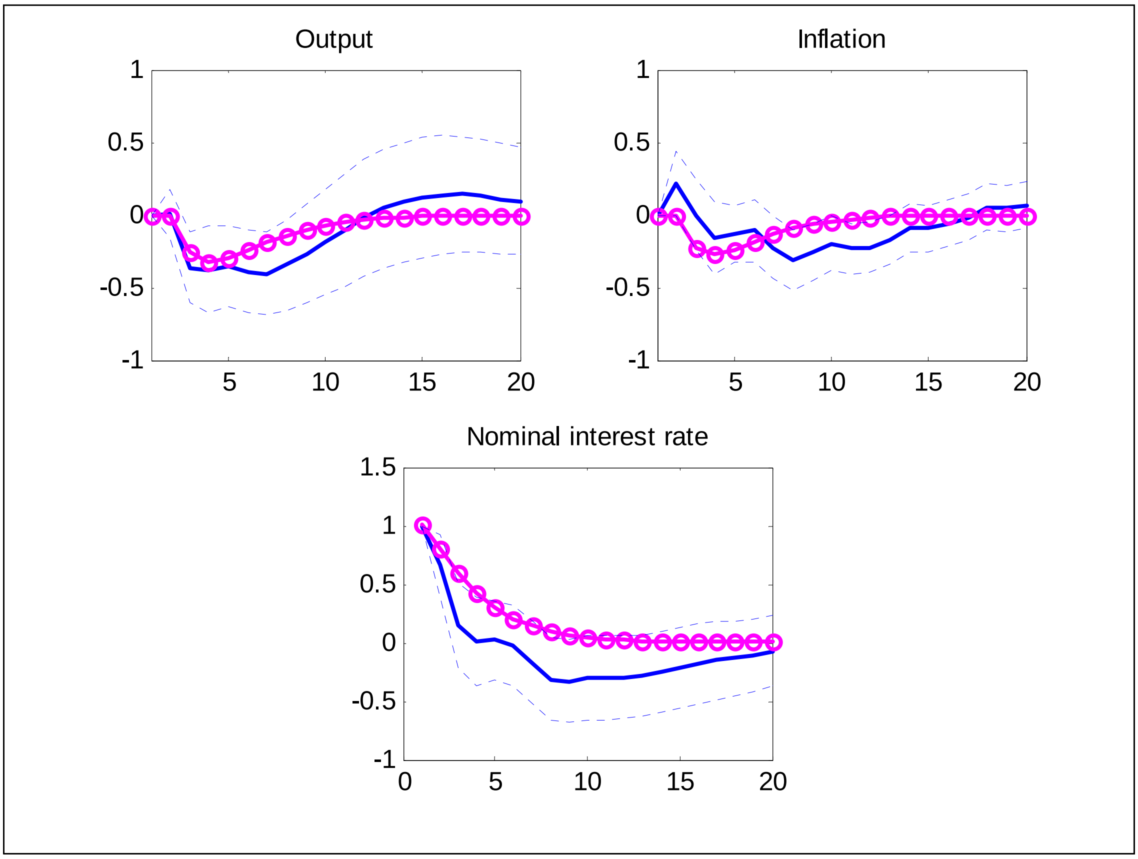

Figure 6: Estimated versus model responses

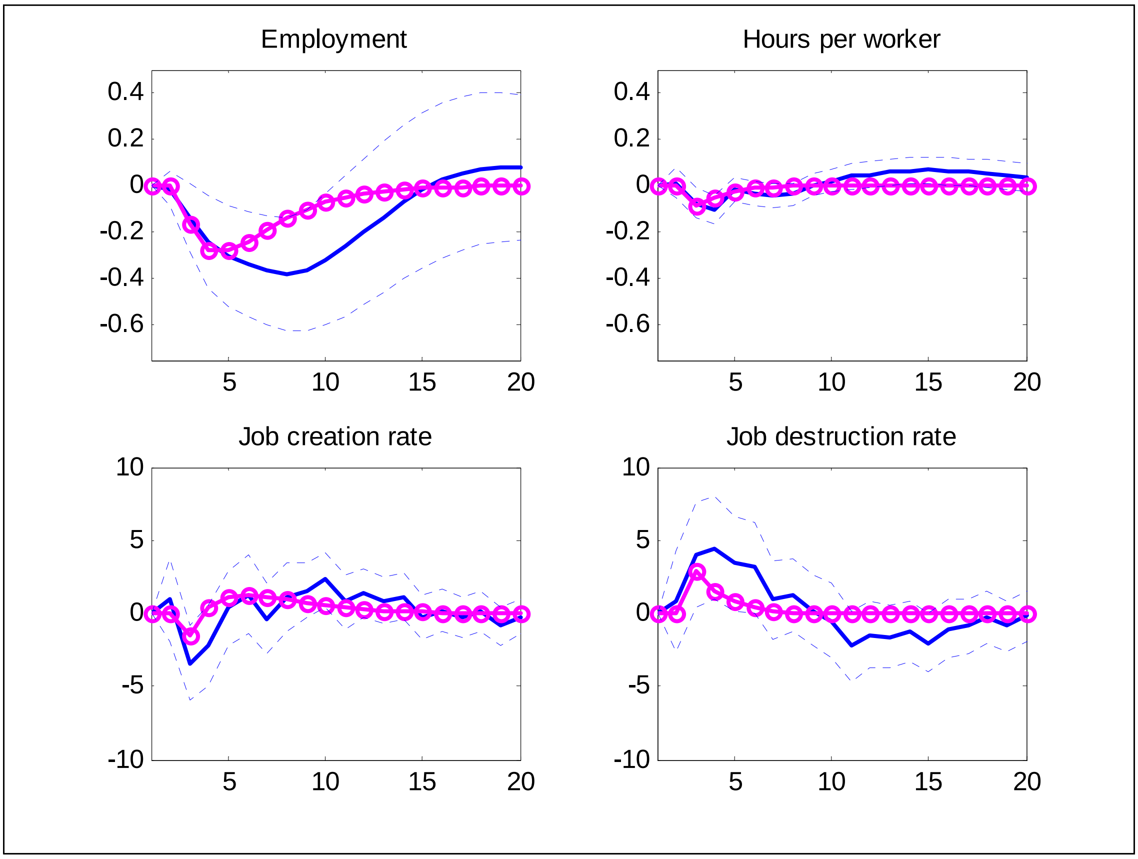

Figure 7: Estimated versus model responses

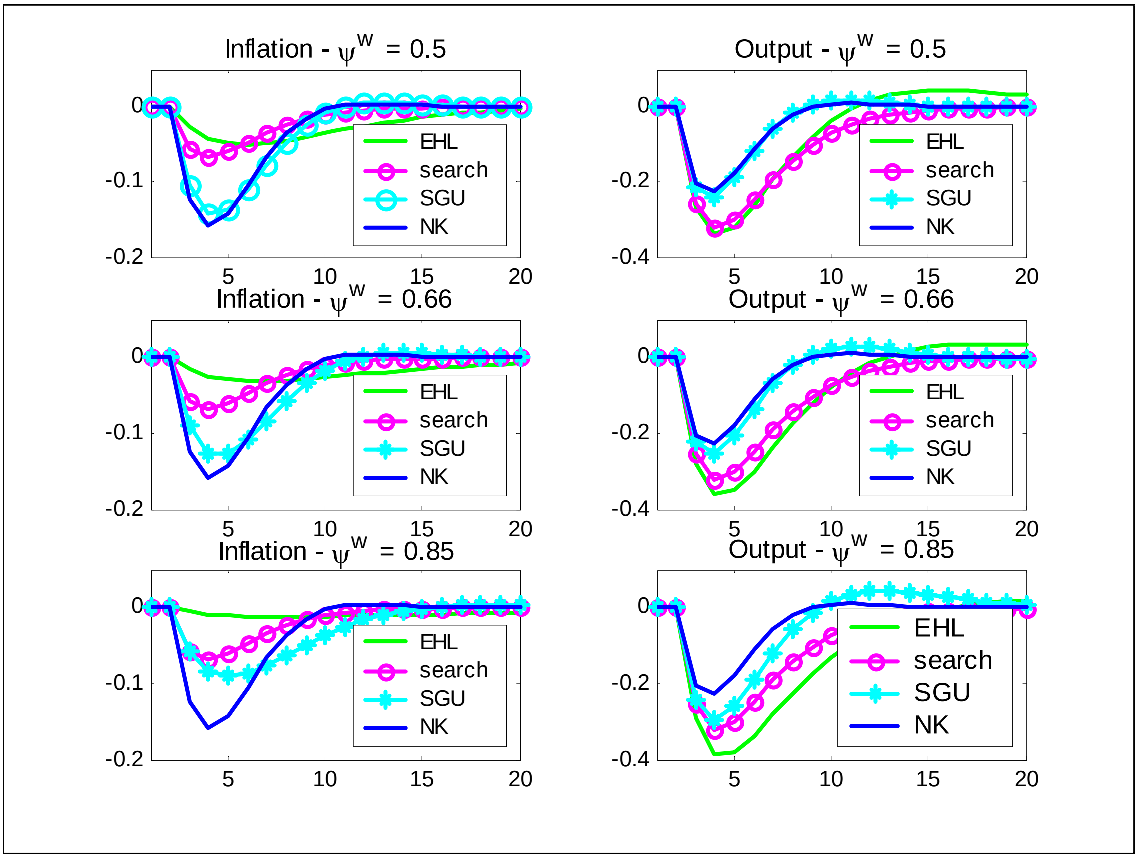

Figure 8: Search versus sticky wage model