

ACCOUNTING FOR THE RELATIONSHIP BETWEEN MONEY AND INTEREST RATES

MAGNUS JONSSON
Sveriges Riksbank

PAUL KLEIN
University of Western Ontario

In time-series from the United States, the relationship between the money to income ratio and the nominal interest rate is a negative and stable one. In Swedish data, there is no such stable relationship. In this paper, we argue that this difference can be explained by the differences in the shock processes that have hit the two countries. Using a dynamic general equilibrium model driven by shock processes estimated to fit the two countries, we find that we can account for the main properties of the data remarkably well.

Keywords: Money, Nominal Interest Rates, Simulated Method of Moments

1. INTRODUCTION

Looking at U.S. data, the relationship between the money-to-income ratio and the nominal interest rate is quite stable and clearly negative. This seems to confirm not just conventional wisdom but also traditional models of money demand [see Cooley and LeRoy (1981) for a good, and critical, survey]. Looking at Swedish data, however, we find that the relationship is weakly positive. Moreover, a survey of 18 different OECD countries exhibits a wide range of patterns (see Figures 1 and 2). Some countries, such as Iceland and Switzerland, are similar to the United States. Others, such as Great Britain and Portugal, look more like Sweden. *Prima facie*, this looks like a challenge to traditional theoretical models.

In this study, we show that this challenge can be met by using the standard tools of the neoclassical tradition. If one takes into account the fact that Swedish and U.S. economies have been hit by different shocks, the basic properties of the data on money and interest rates can be accounted for reasonably well.

We are grateful to Lars Ljungqvist, Lee Ohanian, Torsten Persson, Kjetil Storesletten, Lars E. O. Svensson, and Paul Söderlind, as well as seminar participants at the Federal Reserve Bank of Minneapolis and the European University Institute in Florence for useful discussions and comments. We also are grateful to an anonymous referee for many helpful and constructive comments. The views expressed in this paper are those of the authors and should not be interpreted as the views of Sveriges Riksbank. Address correspondence to: Magnus Jonsson, Research Department, Sveriges Riksbank, 103 37 Stockholm, Sweden; e-mail: magnus.jonsson@riksbank.se.

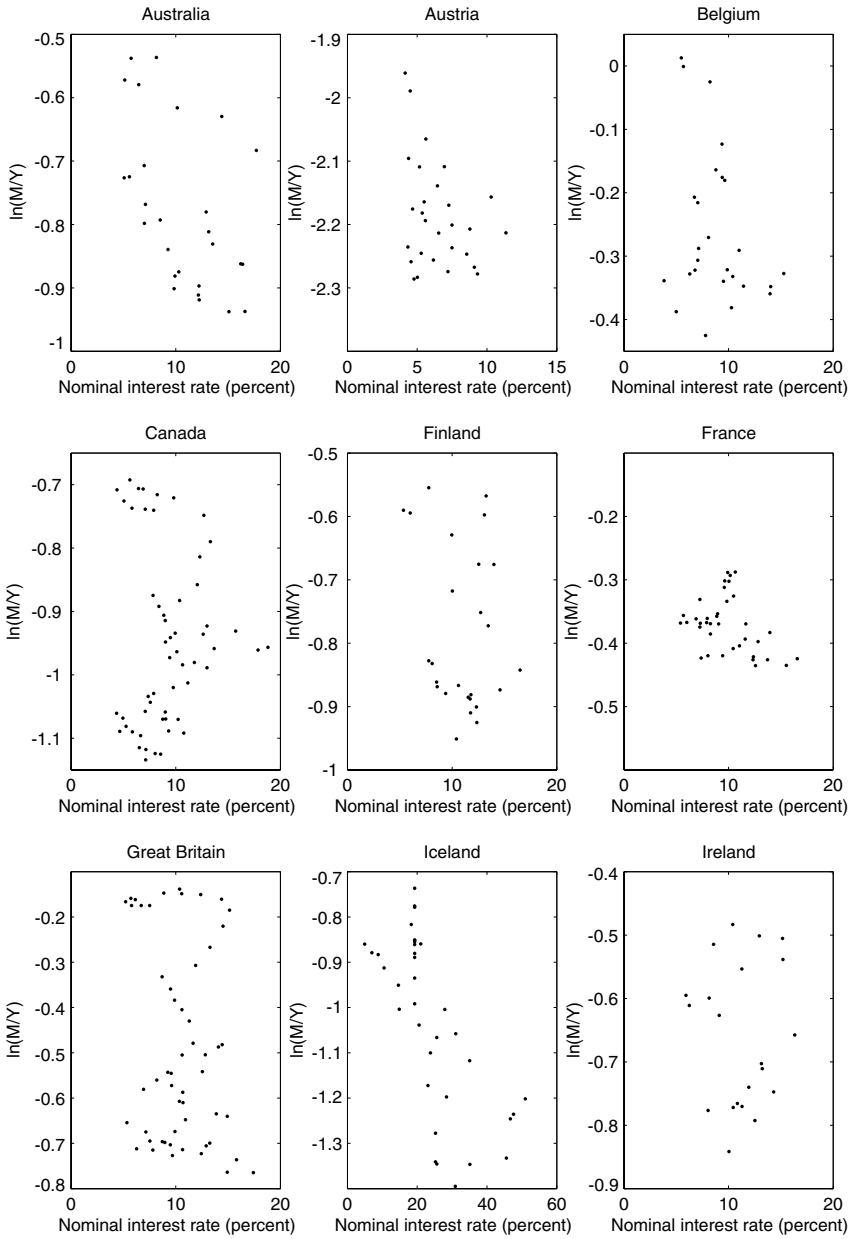


FIGURE 1. $\ln(M/Y)$ versus R in nine OECD countries.

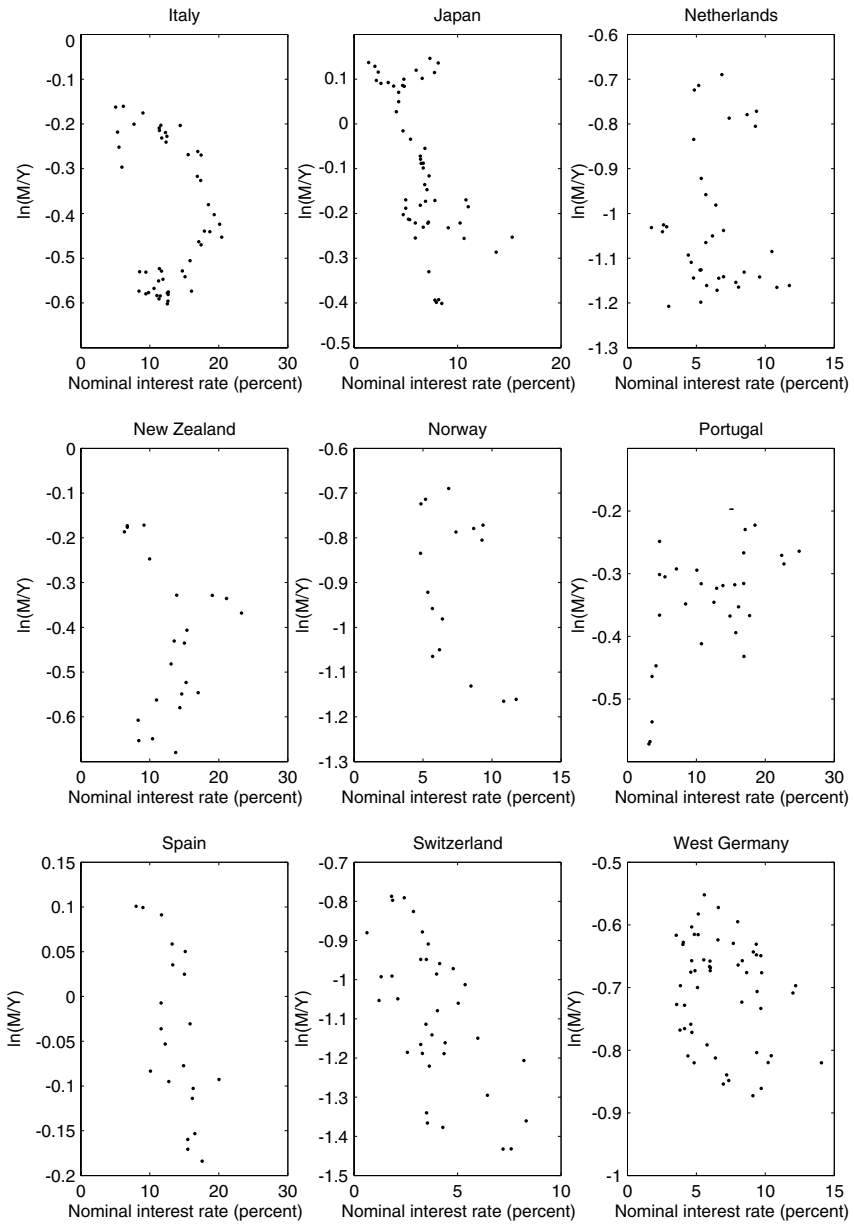


FIGURE 2. $\ln(M/Y)$ versus R in nine other OECD countries.

This result might seem counterintuitive. Indeed, several monetary equilibrium models in the literature, for example, Sidrauski (1967) and Lucas and Stokey (1987), have the property that a deterministic negative relationship between real balances and the nominal interest rate can be derived, and this relationship is independent of the precise nature of the shocks that hit the economy. And even if we would find a model that did not have this property, surely any reasonable model exhibits a negative relation between money and interest rates across steady states. Actually, so do the model we consider here. But comparative steady state relations can be an unreliable guide to the relationships we should expect to find in time series. In the model in this study, the time-series relationship between money and interest rates turns out not to be deterministic but to depend on the relative volatility of the various shocks.

The overall strategy of this study is the following. Treating technology, monetary policy, and fiscal policy as an exogenous vector autoregression (VAR) process, we estimate this process as a VAR using least-squares. Taking this VAR process as given, we fit a monetary general equilibrium model to the data in two different ways. First we take our preference and technology parameters “off-the-shelf”; that is, we use parameter values that have been used in the literature, using the same values for both the United States and Sweden. Second, we formally estimate our model using a combination of the generalized method of moments (GMM) and the simulated method of moments (SMM), a method that also allows us to test the model’s implied moment conditions.

The main result is that the model does well in capturing the differences in the money to income and interest rate relationships in the United States and Sweden. An informal look at the empirical and simulated data reveals that the model is able to reproduce the main features of the relationship. In particular, the model is able to reproduce the differences between the two countries in a striking fashion. That is, the model produces a stable and negative relationship for the United States, whereas it produces a weakly positive relationship for Sweden. Formally, we find that none of our models can be statistically rejected, although we concede that the power of our tests may be low.

We find that the character of this relationship is determined partly by preference and technology parameters but mainly by the differences in the shock processes that have hit the two economies. Technology shocks alone create a near vertical relationship in both countries since those shocks hardly generate any volatility in the nominal interest rate. Adding on money growth rate shocks create a stable and negative relationship in the United States, whereas in Sweden they create only a weak negative relationship. There are two main reasons for this. In the United States, money growth rate shocks have been relatively more persistent than in Sweden. But, more important, money growth rate shocks and technology shocks are positively correlated in the United States, whereas they are almost uncorrelated in Sweden. Whenever money growth is high, technology also will be high. Higher technology increases output, which enhances the fall in the money-to-income ratio. This makes the relationship between the money-to-income ratio and the interest

rate steeper in the United States than in Sweden. The main effect of adding on shocks to fiscal policy is to increase the variance of the interest rate and of the money-to-income ratio in both countries, which brings the relationship closer to what we see in the data.

The study is organized as follows. In Section 2, we present the facts that we would like to explain. Section 3 presents the dynamic general equilibrium model that we use to try to explain these facts. In Section 4, we report the results from the simulations of our model, and Section 5 concludes.

2. THE FACTS

2.1. The United States

Using annual data from 1947 to 1990, the relationship between the natural log of the money to income ratio and the nominal interest rate is depicted in Figure 3. The regression coefficient is -1.95 , the sample standard deviation of the interest rate is 0.030 and the sample standard deviation of the log of the money-to-income ratio is 0.094 .

2.2. Sweden

Using annual data from 1950 to 1990, the relationship between the natural log of the money-to-income ratio and the nominal interest rate is depicted in Figure 3. The regression coefficient is weakly positive; that is, 0.49 with a standard error of 0.31 . Hence, the regression is not significantly different from zero. The sample standard deviation of the interest rate is 0.024 and the sample standard deviation of the log of the money-to-income ratio is 0.049 .

3. THEORETICAL FRAMEWORK

On top of a standard real business cycle model with adjustment costs for capital we impose a shopping time (ST) technology in order to create a role for money. The model is driven by stochastic shocks to technology, government spending, taxes, and the money supply. These shocks are modeled as VAR processes.

The main nonstandard feature of the model is the presence of stochastic fiscal and monetary policy. There are several reasons for including stochastic policy. In the first place, it adds realism in a way that we know is important. That stochastic policy increases the empirical fit of business cycle models has been established for the United States by McGrattan (1994), Braun (1994) and others, and for Sweden by Jonsson and Klein (1996). Also, because the empirical relationship between money and interest rates differs significantly between the United States and Sweden, it makes sense to build into the model those observables that might explain these differences, and the different policy processes are natural candidate explanations. In order to capture the differences between the policy processes to the extent that we need, it seems necessary to include not just

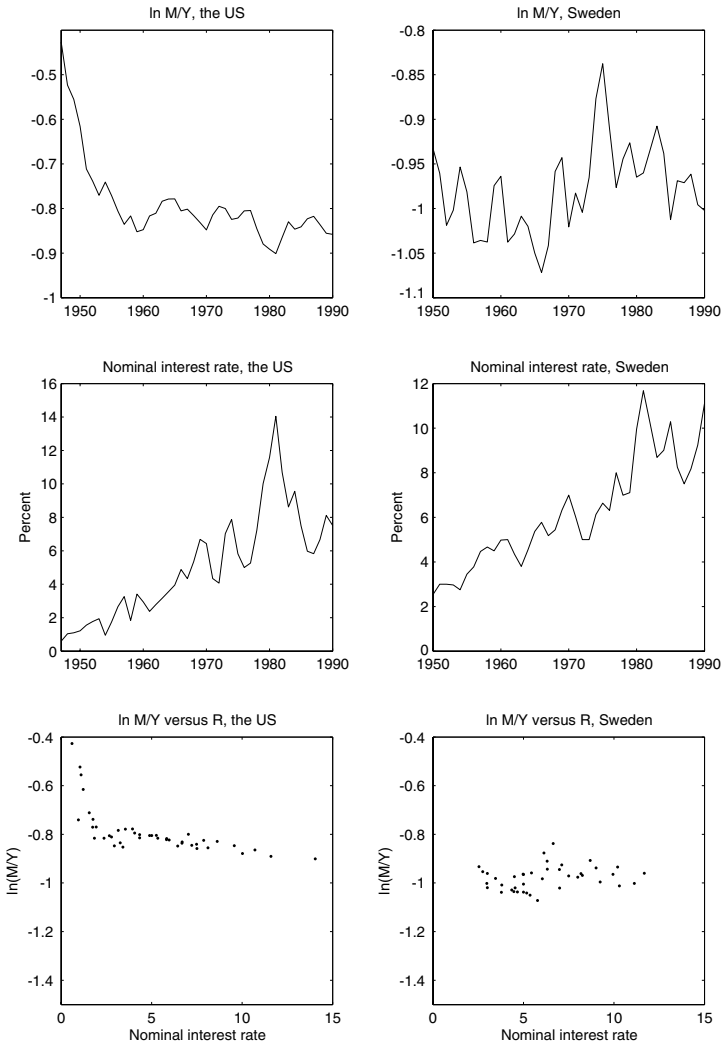


FIGURE 3. Money/income ratios and interest rates in the United States and Sweden.

differences in first moments but in variances and autocorrelations as well. This requires the assumption of stochastic policy, and immediately suggests a (linear) VAR specification because such a specification not only captures variances and autocorrelations well but also easily fits into a linear rational expectation framework.

The tax systems are modeled in slightly different ways for Sweden and the United States. The main reason for this is the availability of data. For example, there exists no good time-series on Swedish tax rates that distinguishes between

labor and capital income taxes. By contrast, we do have good data on the Swedish payroll tax rate, which only affects labor (and not capital) income.

3.1. Households

The economy is inhabited by a large number of households. A representative household has preferences, \mathcal{U} , over stochastic processes for consumption, c , and leisure, ℓ , according to

$$\mathcal{U} = E \left[\sum_{t=0}^{\infty} \beta^t U(c_t, \ell_t) \right], \tag{1}$$

where t denotes time, E the unconditional expectation operator, $U(\cdot)$ the utility function, and $\beta \in (0, 1)$ the subjective discount factor. The period utility function has the standard iso-elastic form

$$U(c_t, \ell_t) = \frac{[c_t^\alpha \ell_t^{1-\alpha}]^{1-\sigma} - 1}{1 - \sigma}, \tag{2}$$

where $\alpha \in (0, 1)$ measures the weight on consumption relative leisure, σ denotes the coefficient of relative risk aversion and $1/\sigma$ the intertemporal elasticity of substitution.

In the model adapted for the United States, the budget constraint, in real terms, is given by

$$\begin{aligned} (1 + \tau_t^c) c_t + i_t + \frac{m_{t+1}}{P_t} + \frac{Q_t b_{t+1}}{P_t} \\ = (1 - \tau_t^k) r_t^k k_t + \tau_t^k \delta k_t + (1 - \tau_t^h) w_t h_t + \frac{m_t}{P_t} + \frac{b_t}{P_t} + T_t, \end{aligned} \tag{3}$$

where m_t denotes the amount of money carried over from period $t - 1$ (and held at the beginning of period t), m_{t+1} the amount of money carried over into period $t + 1$ (and held at the end of period t), Q_t the money price of a nominal bond that delivers one unit of currency in period $t + 1$, b_t the quantity of bonds carried over from period $t - 1$ and that mature in period t , b_{t+1} the quantity of bonds carried over into period $t + 1$, k_t the physical capital stock, i_t the investment expenditure, h_t the fraction of time spent in paid employment, r_t^k the rental rate of capital, w_t the wage rate, P_t the general price level; that is, the money price of goods, τ_t^c the consumption tax rate, τ_t^h the labor income tax rate, τ_t^k the capital income tax rate and δ the depreciation rate, which is modeled as tax deductible. The (real) transfer payments from the government is denoted by T_t , some (but not necessarily all) of which may be in the form of currency. For the model adapted to Sweden, the

corresponding constraint is

$$\begin{aligned}
 & (1 + \tau_t^c) c_t + i_t + \frac{m_{t+1}}{P_t} + \frac{Q_t b_{t+1}}{P_t} \\
 & = (1 - \tau_t^y) (r_t^k k_t + w_t h_t) + \tau_t^y \delta k_t + \frac{m_t}{P_t} + \frac{b_t}{P_t} + T_t,
 \end{aligned} \tag{4}$$

where τ_t^y denotes the general income tax rate. Throughout this paper, capital letters denote aggregated per capita variables and lowercase letters denote individual decision variables. For prices, that is, R_t^k , Q_t , W_t , P_t , capital letters denote nominal prices and lowercase letters denote real prices.

A role for money is introduced through a shopping-time technology, which is a simple way to model the fact that money facilitates transactions. The functional form for the shopping-time technology follows Ljungqvist and Sargent (2000):

$$s_t = \omega_1 \left[\frac{m_{t+1}}{(1 + \tau_t^c) c_t P_t} \right]^{-\omega_2}, \tag{5}$$

where s_t denotes the fraction of time spent “shopping” and ω_1 and ω_2 are parameters. Note that bonds (even when they have matured) cannot be used for shopping. If they could, there would be no role for money in the model. What distinguishes nominal from real bonds is that their real value in terms of goods depends on the general price level in the period of maturity.

To model the idea that capital equipment may be costly to install we introduce adjustment costs for changing the capital stock. The adjustment costs are assumed to be quadratic and zero in steady state and the functional form is taken from Chari, Kehoe, and McGrattan (2000). Hence, capital evolves over time according to

$$k_{t+1} = i_t + (1 - \delta)k_t - \frac{\kappa}{2} \left(\frac{i_t}{k_t} - \delta \right)^2 k_t, \tag{6}$$

where κ quantifies the cost of changing the capital stock.

Finally, the households have a time constraint given by

$$\ell_t + h_t + s_t = 1, \tag{7}$$

where the total time endowment is normalized to one.

3.2. Firms

A representative firm rents capital and employs labor in order to maximize period profits, taking factor prices as given; that is,

$$\max_{\{k_t, h_t\}} [y_t - r_t^k k_t - w_t h_t]. \tag{8}$$

The production technology is Cobb-Douglas, which means that output is produced according to the following formula

$$y_t = z_t k_t^\theta [(1 + \gamma)^t h_t]^{1-\theta}, \tag{9}$$

where γ denotes the real growth rate, θ the capital’s share of output and z_t the technology level.

The rental rate of capital and the wage rate are determined competitively as the marginal products of capital and labor, respectively. For Sweden, we also impose a compulsory social insurance contribution rate, τ_t^w , which drives a wedge between the wage rate and the marginal product of labor in the following way

$$(1 + \tau_t^w) w_t = z_t (1 - \theta) k_t^\theta [(1 + \gamma)^t h_t]^{-\theta}. \tag{10}$$

3.3. Government

The money supply, M_t^s , is determined by the following law of motion

$$M_t^s = u_t M_{t-1}^s \tag{11}$$

where u_t denotes the growth rate of the money supply. The growth rate is in turn determined by the following rule

$$\ln u_t = \ln \bar{u} + \psi_1 \ln \left(\frac{\pi_t}{\bar{\pi}} \right) + \psi_2 \ln \left(\frac{y_t}{y_{t-1}} \right) + \ln \tilde{M}_t^s \tag{12}$$

where \bar{u} denotes the steady state money growth rate, π_t the inflation rate, and $\ln \tilde{M}_t^s$ an exogenously given money supply shock defined as $\tilde{M}_t^s = M_t^s / M_{t-1}^s$. The endogenous part of monetary policy consists of reactions to deviations of inflation from steady state and the growth rate of output. The degree of responsiveness is determined by the parameters $\psi_1 < 0$ and $\psi_2 > 0$. Evidence supportive of monetary policy accommodating output fluctuations can, for example, be found in Finn (1999) and Gavin and Kydland (1999). Ireland (2003) finds empirical evidence for also including reactions to inflation deviations from steady state.

The government’s real budget constraint in the model adapted to the United States is

$$T_t + G_t + \frac{M_{t-1}^s}{P_t} + \frac{B_{t-1}^s}{P_t} = \tau_t^c C_t + \tau_t^h w_t H_t + \tau_t^k (r_t^k - \delta) K_t + \frac{M_t^s}{P_t} + \frac{Q_t B_t^s}{P_t}, \tag{13}$$

where G_t denotes government consumption and B_t^s the quantity of bonds supplied by the government at the end of period t , whereas for the model adapted for

Sweden it is

$$\begin{aligned}
 T_t + G_t + \frac{M_{t-1}^s}{P_t} + \frac{B_{t-1}^s}{P_t} \\
 = \tau_t^c C_t + \tau_t^w w_t H_t + \tau_t^y [(r_t^k - \delta) K_t + w_t H_t] + \frac{M_t^s}{P_t} + \frac{Q_t B_t^s}{P_t}.
 \end{aligned}
 \tag{14}$$

This implies that T_t is determined residually (and hence endogenously), whereas all other government policy variables are exogenous.

In addition to the notation introduced earlier, we will use $R_t = 1/Q_t - 1$ to denote the one-period nominal interest rate.

3.4. The Exogenous Shock Processes

We will consider three different cases; in the first, there are shocks only to technology, in the second to technology and money growth, and in the final case we allow for shocks to fiscal policy as well. In each case, we perform a separate estimation of the resulting VAR-model. However, the mean of the shocks is fixed to the sample mean in each of the three cases. The following VAR-model is estimated

$$X_t = (I - \varphi) \mu + \varphi X_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma),
 \tag{15}$$

In the final case with all the shocks, X_t consists of the following variables for the United States and Sweden:¹

$$X_t^{US} = \begin{bmatrix} \ln z_t \\ \ln g_t \\ \tau_t^c \\ \tau_t^h \\ \tau_t^k \\ \ln \tilde{M}_t^s \end{bmatrix}, \quad X_t^{SWE} = \begin{bmatrix} \ln z_t \\ \ln g_t \\ \tau_t^c \\ \tau_t^w \\ \tau_t^y \\ \ln \tilde{M}_t^s \end{bmatrix}.
 \tag{16}$$

The parameters μ , φ , and Σ are estimated using least-squares. For simulation purposes, we need to construct disturbance vectors with the estimated Σ as their variance matrix. For that purpose, we multiply standard normal vectors by the matrix Ψ where Ψ is the lower triangular Cholesky decomposite of the estimated Σ matrix.

For purposes of interpretation, it is worth considering the unconditional variances and unconditional autocorrelations of the elements of the two estimated shock processes (see Table 1). Overall, we see that the Swedish economy has been hit by more volatile shocks than has the United States, and that monetary and fiscal policy has been much more volatile relative to technology in Sweden than in the United States.

TABLE 1. Sample moments of the shocks in the United States and Sweden

US				Sweden			
Variable	Mean	std dev	ac	Variable	Mean	std dev	ac
$\ln z_t$	0.000	0.0604	0.84	$\ln z_t$	0.000	0.0649	0.88
$\ln g_t$	-1.658	0.1282	0.69	$\ln g_t$	-1.294	0.2289	0.92
τ_t^c	0.057	0.0061	0.90	τ_t^c	0.239	0.0703	0.90
τ_t^h	0.235	0.0334	0.94	τ_t^w	0.204	0.1354	0.95
τ_t^k	0.514	0.0448	0.76	τ_t^y	0.488	0.0714	0.85
$\ln \tilde{M}_t^s$	0.050	0.0299	0.58	$\ln \tilde{M}_t^s$	0.085	0.0456	0.36

Notes: Std dev denotes the standard deviation and ac the autocorrelation.

3.5. Equilibrium and Solution Algorithm

We assume a competitive equilibrium in which behavior is identical across households and across firms. This allows us to treat the economy as comprising of a representative household and a representative firm. An equilibrium consists of prices and quantities, such that:

1. Each household chooses consumption, investment, capital, leisure, labor, bonds, and money holdings to maximize its expected lifetime utility subject to its constraints and given government policy.
2. Each firm chooses capital and labor to maximize profits subject to its constraints and given government policy.
3. The government satisfies its budget constraint.
4. In addition to aggregate consistency, the aggregate resource constraint hold and markets clear; that is,

$$Y_t = C_t + I_t + G_t + \frac{\kappa}{2} \left(\frac{i_t}{k_t} - \delta \right)^2 k_t, \tag{17}$$

$$M_t^s = M_{t+1}, \tag{18}$$

$$B_t^s = B_{t+1} = 0. \tag{19}$$

To ensure the existence of a time-invariant decision rule, all variables must be stationary. For the real variables, we assume that technological growth follows a geometric trend at rate γ and proceed in the manner of Hansen and Prescott (1995).² Nominal variables are detrended with the nominal price; that is, real money balances are assumed to be stationary. To solve for the decision rules, we linearize the stationary first-order conditions around steady state values and solve for the saddle path of the resulting linear system of difference equations, using the algorithm proposed in Klein (2000).

4. SIMULATION RESULTS

How to choose the technology and preference parameters of business cycle models has been the subject of some controversy [see Hansen and Heckman (1996) and

TABLE 2. Parameter values

Parameter	Off-the-shelf		SMM	
	US	SWE	US	SWE
α	0.38	0.38	0.38	0.38
β	0.95	0.95	0.95	0.95
γ	0.02	0.02	0.02	0.02
δ	0.10	0.10	0.10	0.10
θ	0.36	0.36	0.36	0.36
κ	2.00	2.00	0.00	3.90
σ	1.00	1.00	1.00	1.00
ψ_1	-0.50	-0.50	-0.55	-1.05
ψ_2	0.50	0.50	0.85	0.36
ω_1	0.006	0.006	0.004	0.001
ω_2	2.00	2.00	2.52	3.77

Kydland and Prescott (1996)]. In this study, we identify and implement two different ways of choosing parameters.

The first method is to take preference and technology parameters “off-the-shelf”; that is, to use the same parameters as others in the literature have done. The main advantage of this approach is that it enables us to identify the source of the different predictions for Sweden and the United States as the differences in the shock processes rather than differences in the technology and preference parameters. This means that our explanation of the facts is in terms of observable shocks rather than unobservable parameters.

The second method is to estimate some of the parameters separately for each country using a combination of GMM and SMM. One advantage with this method is that, if the weighting matrix is chosen optimally, less weight is given to uncertain moments. Also, if there are more moments than parameters to be estimated, the method provides us with a way of formally testing the hypothesis that the model’s moments are equal to the corresponding empirical moments.

4.1. Off-the-Shelf Parameters

The parameter values are summarized in Table 2. Many of these values can be considered as standard since they have appeared in many studies, see for example Cooley and Prescott (1995) and Christiano, Eichenbaum, and Evans (2005). In particular, the values for α , β , γ , δ , θ , and σ are simply taken from these studies. This imply, among other things, that we set $\sigma = 1$; that is, we are assuming log-utility. The real growth rate, γ , is set equal to 0.02, which is close to the average growth rate in both countries over the period. For the capital adjustment cost parameter, κ , the parameters in the shopping time technology, ω_1 and ω_2 , and the parameters in the monetary policy rule, ψ_1 and ψ_2 , it is not possible to pick standard values from the literature because these values varies between different models. We choose them in the following way: κ , ω_1 and ω_2 , are adjusted

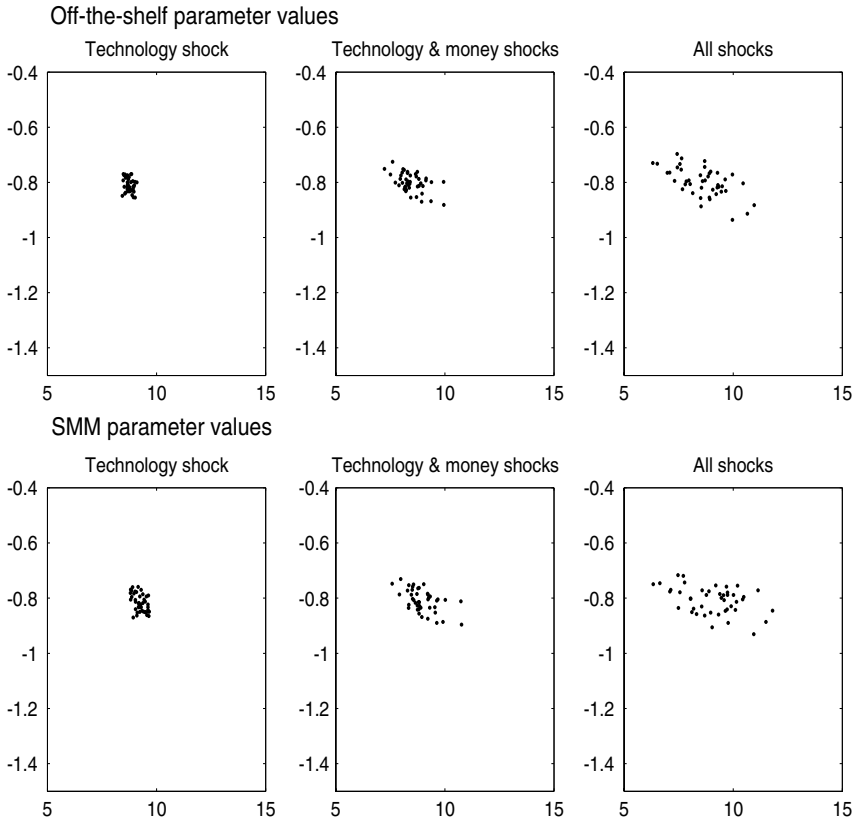


FIGURE 4. Simulated results, the United States.

simultaneously in order to match the following three moments

$$\frac{\sigma \ln I_t}{\sigma \ln Y_t}, \quad \mu \ln(M_t/Y_t), \quad \sigma \ln(M_t/Y_t), \quad (20)$$

where σ denotes the standard deviation and μ the mean (abusing the notation somewhat). It is not possible to exactly match these moments to the data because we restrict the parameters to be the same for both countries. Nevertheless, we succeed in matching these moments fairly well in both countries except for the standard deviation of the money-to-income ratio in the United States, which is somewhat too low in the model compared to the data. The values for the parameters in the money supply rule are taken from the estimates given in Jonsson and Palmqvist (2003). This means that the value for ψ_2 is well in the range of plausible values considered by Finn (1999).

Turning now to the results for the United States, Figure 4 shows the simulated relationship between the log of the money-to-income ratio and the nominal interest rate under different shock processes. We first consider only technology shocks.

TABLE 3. Simulated and empirical moments in the United States

Moment	Off-the-shelf		SMM		Empirical 1947–90
	Value	[90% conf. int.]	Value	[90% conf. int.]	
$\beta_{\ln(M_t/Y_t), R_t}$	-3.21 (0.83)	[-4.24 -1.50]	-2.33 (1.00)	[-3.72 -0.46]	-1.95 (0.34)
$\rho_{\ln(M_t/Y_t), R_t}$	-0.61 (0.16)	[0.087 0.105]	-0.49 (0.20)	[-0.72 -0.09]	-0.65 (0.08)
$\rho_{\ln(M_t/P_t), R_t}$	-0.27 (0.30)	[-0.71 0.27]	-0.33 (0.29)	[-0.75 0.19]	-0.16 (0.29)
$\rho_{\ln(M_t/P_t), Y_t}$	0.61 (0.22)	[0.15 0.84]	0.65 (0.22)	[0.17 0.86]	0.44 (0.29)
σ_{R_t}	0.009 (0.002)	[0.006 0.012]	0.010 (0.002)	[0.007 0.014]	0.030 (0.005)
$\sigma_{\ln(M_t/Y_t)}$	0.048 (0.009)	[0.034 0.062]	0.050 (0.010)	[0.034 0.066]	0.094 (0.030)
σ_{π_t}	0.026 (0.003)	[0.021 0.031]	0.023 (0.003)	[0.018 0.027]	0.034 (0.005)
μ_{R_t}	0.085 (0.004)	[0.079 0.092]	0.085 (0.004)	[0.082 0.096]	0.049 (0.010)
$\mu_{\ln(M_t/Y_t)}$	-0.79 (0.02)	[-0.82 -0.76]	-0.79 (0.02)	[-0.82 -0.77]	-0.79 (0.03)
$\chi^2(2)$	NA		0.29		
p -value	NA		0.87		

Notes: In this table, μ denotes the mean, σ the standard deviation, ρ the correlation coefficient, and β the regression coefficient, abusing the notation somewhat. Standard errors are in parentheses. A 90% simulated confidence interval are in brackets. Statistics for the model economy are computed on data generated by simulating the model for 44 periods and repeating the simulation 1,000 times (each sample is initiated with 200 extra observations, which are then discarded). The various simulated moments are then means over these 1,000 simulations, and the standard errors are the standard deviations. The confidence intervals are also computed from the simulations. The standard errors of the empirical moments are computed either in a standard way, or, in the case of the correlation coefficients, by combining the method of Newey and West (1987) with the delta method.

In this case the variability of the interest rate almost vanishes, which induces a near-vertical relationship between the money to income ratio and the interest rate. Adding on monetary shocks, the simulated regression coefficient becomes negative and the variance of the interest rate increases. In the final case, where we also add on shocks to fiscal policy, the variance of both the interest rate and the money-to-income ratio increases. Hence, bringing the relationship closer to what we see in the data.

Table 3 reports a number simulated moments related to the money-to-income ratio and the interest rate together with their empirical counterparts. All shocks are included in the VAR-model when we calculate the simulated moments. The table also reports standard errors and a simulated 90% confidence interval. The simulated regression coefficient is -3.21 , whereas it is -1.95 in the data. The standard error is fairly high, which implies that the empirical value is within the 90% simulated confidence interval. The model is also close to replicate the correlation coefficient: the empirical value is -0.61 , whereas the simulated value is -0.79 .

The overall picture is that the model does a fairly good job in replicating the other moments as well. However, the model fails to replicate the standard deviation of

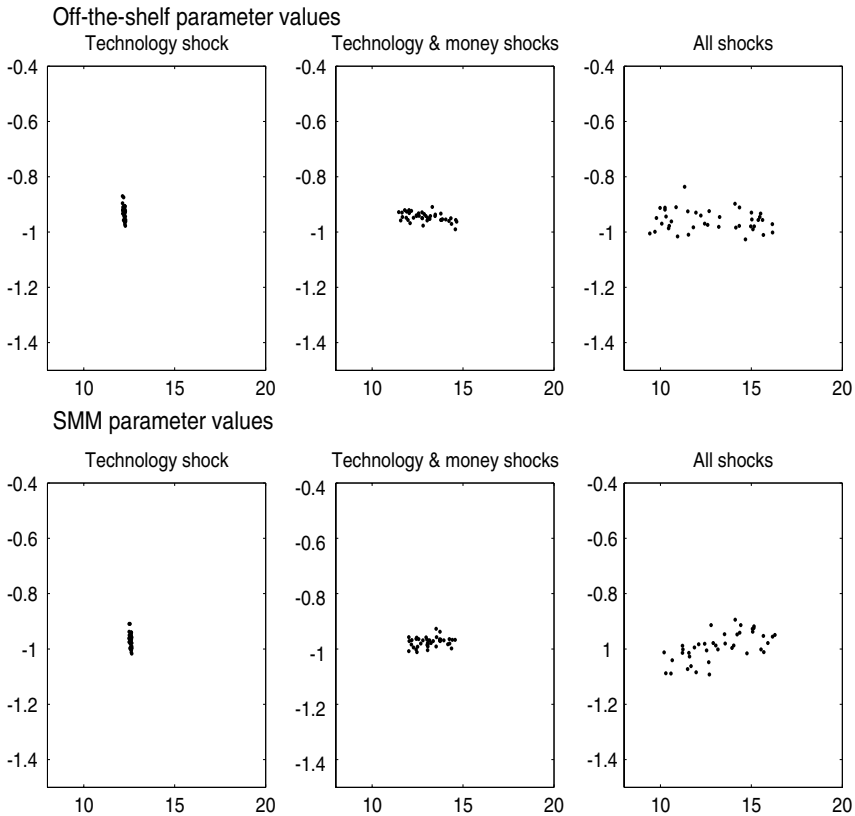


FIGURE 5. Simulated results, Sweden.

some of the variables. In particular, we have difficulties in reproducing the volatility of the interest rate relative to that of the inflation rate, and this problem seems unfortunately to plague this class of models generally [see den Haan (1995)].³

The results for the Swedish economy are found in Figure 5 and Table 4. Technology shocks give rise to a similar relationship as in the United States. In contrast to the U.S. case, adding on shocks to the money growth rate only creates a weak negative relationship between the money-to-income ratio and the interest rate. When we add on shocks to fiscal policy, these shocks increase the variance of both the money-to-income ratio and the interest rate, whereas they have a small effect on the slope.

The simulated regression coefficient is -1.14 with a 90% simulated confidence interval between -2.33 and 0.77 . The empirical value of 0.49 is thus within this interval. Regarding the other moments reported in Table 4, it is notable that the model is close to reproduce the standard deviation of the money-to-income ratio and that the standard deviation of the interest rate is within the simulated

TABLE 4. Simulated and empirical moments in Sweden

Moment	Off-the-shelf		SMM		Empirical 1950–90
	Value	[90% conf. int.]	Value	[90% conf. int.]	
$\beta_{\ln(M_t/Y_t), R_t}$	-1.14 (0.96)	[-2.33 0.77]	0.26 (0.87)	[-0.93 1.84]	0.49 (0.31)
$\rho_{\ln(M_t/Y_t), R_t}$	-0.37 (0.27)	[-0.70 0.17]	0.09 (0.24)	[-0.31 0.51]	0.24 (0.10)
$\rho_{\ln(M_t/P_t), R_t}$	0.15 (0.32)	[-0.45 0.61]	0.11 (0.30)	[-0.43 0.55]	-0.08 (0.23)
$\rho_{\ln(M_t/P_t), Y_t}$	0.94 (0.10)	[0.69 0.98]	0.96 (0.08)	[0.76 0.99]	0.71 (0.09)
σ_{R_t}	0.015 (0.005)	[0.007 0.024]	0.014 (0.004)	[0.008 0.022]	0.024 (0.003)
$\sigma_{\ln(M_t/Y_t)}$	0.047 (0.008)	[0.034 0.061]	0.041 (0.006)	[0.032 0.052]	0.049 (0.006)
σ_{π_t}	0.037 (0.005)	[0.030 0.046]	0.029 (0.004)	[0.018 0.027]	0.033 (0.004)
μ_{R_t}	0.123 (0.005)	[0.114 0.132]	0.123 (0.005)	[0.119 0.134]	0.060 (0.008)
$\mu_{\ln(M_t/Y_t)}$	-0.95 (0.03)	[-1.00 - 0.90]	-0.99 (0.03)	[-1.03 - 0.94]	-0.98 (0.01)
$\chi^2(2)$	NA		5.25		
p-value	NA		0.07		

Notes: In this table, μ denotes the mean, σ the standard deviation, ρ the correlation coefficient, and β the regression coefficient, abusing the notation somewhat. Standard errors are in parentheses. A 90% simulated confidence interval are in brackets. Statistics for the model economy are computed on data generated by simulating the model for 41 periods and repeating the simulation 1,000 times (each sample is initiated with 200 extra observations which are then discarded). The various simulated moments are then means over these 1,000 simulations, and the standard errors are the standard deviations. The confidence intervals are also computed from the simulations. The standard errors of the empirical moments are computed either in a standard way, or, in the case of the correlation coefficients, by combining the method of Newey and West (1987) with the delta method.

confidence interval.⁴ We conclude that our model driven by different shock processes for the United States and Sweden perform fairly well in reproducing the empirical relationship between the money-to-income ratio and the interest rate in both countries as well as a number of other related moments.

4.2. SMM Parameters

In this subsection, we formally estimate the model using a combination of GMM and SMM (see Appendix B for a formal description). The parameters we estimate are the adjustment cost parameter, the parameters in the shopping-time technology, and the parameters in the monetary policy rule; that is, the following five parameters: κ , ω_1 , ω_2 , ψ_1 , and ψ_2 . For the other parameters, that is, α , β , γ , δ , θ , and σ , we use the same values as in the off-the-shelf case, as these values can be considered as standard.⁵

The parameter estimates are reported in Table 2. The values are somewhat different compared to the off-the-shelf case, for example, the capital adjustment cost parameter is now zero for the United States, whereas it is 3.90 for Sweden. Overall, the model’s moment conditions cannot be rejected neither for the United States nor Sweden (see Tables 3 and 4).

For details of the results, see Tables 3 and 4 as well as Figures 4 and 5. The results are quite similar to the case with off-the-shelf parameters. However, the overall fit of the model is better for both countries. This is because we now choose some of the parameters in order to make the model replicate the empirical moments. For example, the regression coefficient is closer to the data, both in the United States and in Sweden. In the Swedish case, the coefficient is in fact weakly positive as in the data. The confidence interval is still large, however. The reason the model does not exactly replicate the regression coefficients is because of the uncertainty associated with this moment, which leads to a low weight being given to it in the estimation.

4.3. How to Understand the Results

We report the response of the money-to-income ratio and the nominal interest rate to a one standard-deviation shock to the exogenous variables (see Figures 6 and 7). This enables us to better understand the mechanisms behind the results. The response of the nominal interest rate is shown with a solid line and the money-to-income ratio is shown with a dashed line.

We have seen that technology shocks give rise to a near-vertical relationship between the money-to-income ratio and the interest rate. The impulse response to a positive technology shock illustrates this finding. Increased technology induces a negligible effect on the interest rate.⁶ At the same time, the shock leads to quantitatively large effects on the money to income ratio. It falls in the first period and rises in the second. This can be understood in the following way: Higher productivity means higher output as well as higher demand for real money balances in order to facilitate “shopping.” However, real money balances are predetermined, which implies that in the period of the shock they are not allowed to change. The money-to-income ratio then falls in the first period as a result of higher output. In the next period, this ratio increases when households choose to hold more real balances.

Shocks to the money growth rate imply a negative relationship between the money-to-income ratio and the interest rate. The money growth rate is persistent, so a positive shock to it tends to increase inflation expectations and hence the nominal interest rate. And because the nominal interest rate represents the opportunity cost of holding cash, there is a tendency for agents to substitute out of money when the interest rate is high and into it when it is low. Hence, when the nominal interest rate is high because of a money growth shock, the ratio of money-to-income falls. This mechanism leads to a negative relationship between the money-to-income ratio and the nominal interest rate when the model economies are exposed to money growth shocks.

We have seen that adding shocks to the money growth rate imply a steeper slope in the United States than in Sweden. To understand this result, it is necessary to note how the money growth rate shock is correlated with the other exogenous shocks. In the U.S. data, shocks to the money growth rate and technology are positively correlated, whereas they are almost uncorrelated in the Swedish

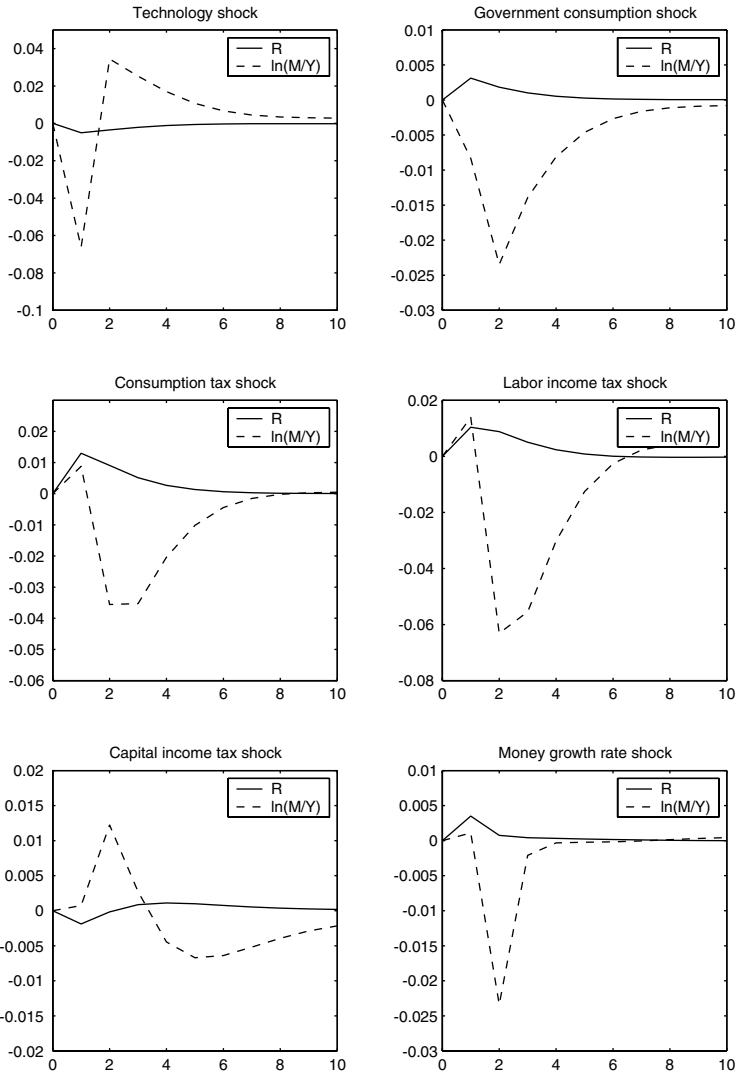


FIGURE 6. Impulse responses, the United States.

data. This means that when money growth is high in the United States, technology also is high. As a consequence of higher technology, output increases, which decreases the money-to-income ratio further. Hence, the slope will be steeper with U.S. shocks than with Swedish shocks. Another but quantitatively less important reason is that shocks to the money growth rate are more persistent in the United States than in Sweden (see Table 1).

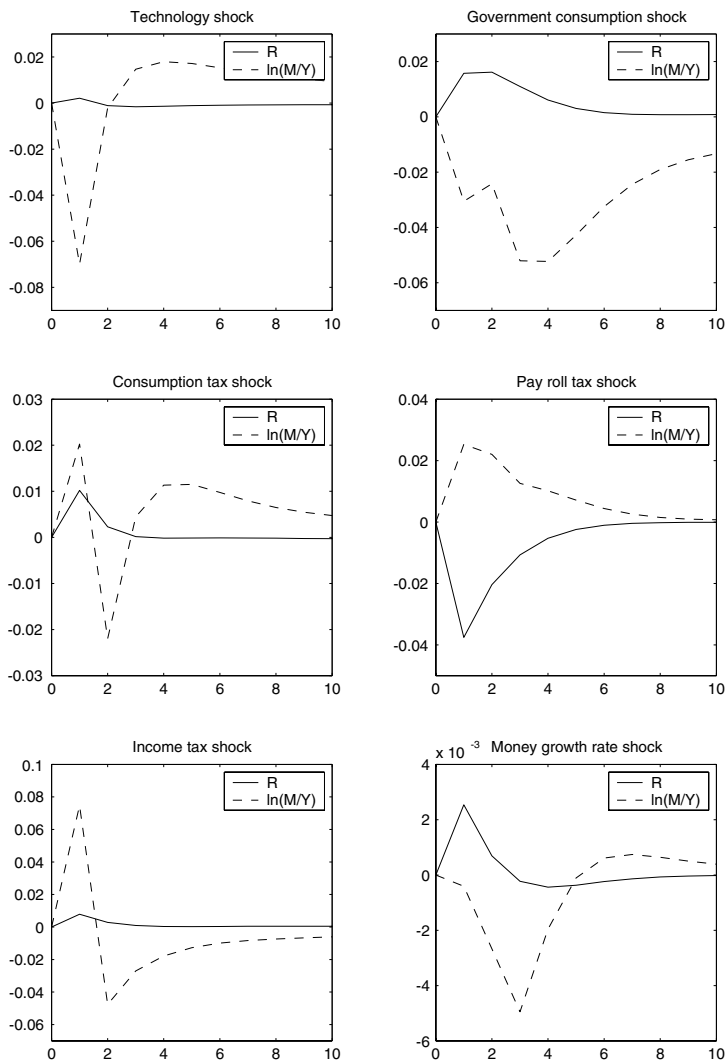


FIGURE 7. Impulse responses, Sweden.

Regarding fiscal policy, it is tempting to argue that its effect is ambiguous, leading to a more chaotic relationship whenever it is volatile. This is to some extent also borne out by the impulse responses. However, shocks to government consumption are an exception. They unambiguously lead to a negative relationship. The reason is the following. An increase in government consumption leads to lower private consumption. This causes the price of real government bonds to fall, which also leads to cheaper nominal bonds. This implies a higher nominal interest rate. Because the nominal interest rate is the opportunity cost of holding cash,

real money balances falls. Output initially rises as a result of higher government consumption, but because both private consumption and investment falls, output also falls in the following periods.⁷ The fall in output mitigates the fall in the money-to-income ratio somewhat.

In general, shocks to the tax rates lead to weak correlations between the money-to-income ratio and the interest rate. In particular, the effect depends on how shocks to the tax rates are correlated with other shocks in the VAR-model. That is why, for example, a shock to the consumption tax rate looks quite different in the United States compared to Sweden. It turns out that the main quantitative effect of shocks to the tax rates is to increase the variance in both the money to income ratio and the interest rate. Because tax rates are more volatile in Sweden than in the United States, this effect is more pronounced in Sweden. The impulse responses still suggest that, in the Swedish case, shocks to the consumption and income tax rates induce a weak positive correlation. For the United States, by contrast, shocks to each of the taxes on their own imply a weak negative relationship between the money-to-income ratio and the nominal interest rate.

In summary: The reason the model predicts a stable and negative relationship between the money-to-income ratio and the interest rate in the United States is that (i) it has been hit by shocks to the money growth rate that are positively correlated with technology shocks, which gives rise to a clear negative relationship; and (ii) it has been hit by shocks to the tax rates that induce a small negative impact on the relationship. The reason the model predicts no relationship or a weakly positive one in Sweden is that (i) shocks to the money growth rate induces only a weak negative relationship and (ii) shocks to the consumption and income tax rates lead to a weakly positive relationship.

5. CONCLUSIONS

In the United States the relationship between the money-to-income ratio and the nominal interest rate is a negative and stable one, while in Sweden there is no such stable relationship. We have argued that the basic neoclassical growth model is able to account for the differences between Sweden and the United States if we take into account that the two countries have been hit by different shocks. The pattern has been clear. Technology shocks alone do not generate enough variance in the nominal interest rate to create much of a relationship at all. Money growth rate shocks, on the other hand, can, and tend to create a stable and negative relationship. And we have seen that the magnitude of the slope is smaller in the model driven by Swedish shocks than in those driven by U.S. shocks. Finally, shocks to fiscal policy tend to increase the volatility of both the money-to-income ratio and the interest rate, which brings the models closer to the data.

This paper has focused on differences in the exogenous shock processes in the two countries; that is, technology, monetary and fiscal policy shocks. We have thus neglected other explanations that could be of interest. For example, Sweden is a small open economy and has therefore also been subject to foreign shocks. We have

not formally modelled any credit market frictions, although the capital adjustment cost could be interpreted as a reduced form for that. Still, extending the model with a formal credit market is presumably an interesting area for future research. The framework developed by Carlstrom and Fuerst (1997) may be useful for this purpose. Other frictions, like price and wage stickiness, could also be interesting to study since they may have been different in Sweden and the United States.

NOTES

1. Since government consumption is nonstationary, we use the ratio of government consumption to GDP, which is denoted by g_t .

2. This procedure implies that the effective discount factor is not β but $(1 + \gamma)^{\alpha(1-\sigma)} \beta$. Abusing the notation somewhat, the letter β will be used below to denote this effective discount factor.

3. See also Hornstein and Uhlig (2000) for a possible way to overcome this problem.

4. Note that the model does not replicate the average nominal interest rate of about 6 percent in the data. The reason is the following: the money growth rate has been almost 9 percent in the data, which implies an average inflation rate of almost 7 percent in our model (after deducting the real growth rate). This implies that the average real interest needs to be -1 percent to replicate the nominal interest rate; that is, the subjective discount factor needs to be larger than 1. We find this implausible and therefore neglect this apparent puzzle in the Swedish data. Hence, we use the standard value for the discount rate.

5. By construction, all our simulated series are stationary. Some of the empirical ones are not, however, because of real secular growth. Since the calculation of first and second moments requires stationarity in order for them to be well-defined, we filter some of the variables (both empirical and simulated) before estimation. In particular, a linear trend is subtracted from $\ln Y_t$ and $\ln (M_t/P_t)$. The remaining series are left unfiltered.

6. This is because investment responds to the technology shock and smooths the interest rate through diminishing returns to capital.

7. The effect of a government consumption shock on investment is ambiguous in general, but in our model the effect turns out to be negative.

REFERENCES

- Braun, Anton (1994) Tax disturbances and real economic activity in the postwar United States. *Journal of Monetary Economics* 33, 441–62.
- Carlstrom Charles and Timothy Fuerst (1997) Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review* 87, 893–910.
- Chari, V.V., Patrick Kehoe, and Ellen R. McGrattan (2000) Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem. *Econometrica* 68, 1151–1180.
- Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005) Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113, 1–45.
- Cooley, Thomas F. and Stephen LeRoy (1981) Identification and estimation of money demand. *American Economic Review* 71, 825–44.
- Cooley, Thomas F. and Edward C. Prescott (1995) Economic growth and business cycles. In Thomas F. Cooley (ed.), *Frontiers of Business Cycle Research*, pp. 1–38. Princeton, N.J.: Princeton University Press.
- Finn, Mary (1999) An equilibrium theory of nominal and real exchange rate comovement. *Journal of Monetary Economics* 44, 453–475.
- Friedman, Milton and Anna Schwartz (1982) *Monetary Trends in the United States and the United Kingdom*. Chicago: University of Chicago Press.

- Gavin, William and Finn E. Kydland (1999) Endogenous money supply and the business cycle. *Review of Economic Dynamics* 2, 347–369.
- den Haan, Wouter (1995) The term structure of interest rates in real and monetary economies. *Journal of Economic Dynamics and Control* 19, 909–40.
- Hansen, Gary D. and Edward C. Prescott (1995) Recursive methods for computing equilibria of business cycle models. In Thomas F. Cooley (ed.), *Frontiers of Business Cycle Research*, pp. 39–64. Princeton N.J.: Princeton University Press.
- Hansen, Lars Peter and James Heckman (1996) The empirical foundations of calibration. *Journal of Economic Perspectives* 10, 87–104.
- Hornstein, Andreas and Harald Uhlig (2000) What is the real story for interest rate volatility? *German Economic Review* 1, 43–67.
- Ireland, Peter (2003) Endogenous money or sticky prices? *Journal of Monetary Economics* 50, 1623–1648.
- Jonsson, Gunnar and Paul Klein (1996) Stochastic fiscal policy and the Swedish business cycle. *Journal of Monetary Economics* 38, 245–268.
- Jonsson, Magnus and Stefan Palmqvist (2003) Inflation, Markups, and Monetary Policy. Sveriges Riksbank Working Paper Series No. 148.
- Klein, Paul (1995) Swedish Income Taxes 1951–1990. Manuscript, IIES, Stockholm University.
- Klein, Paul (2000) Using the generalized Schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control* 24, 1405–1423.
- Kydland, Finn E. and Edward C. Prescott (1991) Hours and employment variation in business cycle theory. *Economic Theory* 1, 63–81.
- Kydland, Finn E. and Edward C. Prescott (1996) The computational experiment: An econometric tool. *Journal of Economic Perspectives* 10, 69–85.
- Lee, Bong Soo and Beth F. Ingram (1991) Simulation estimation of time-series models. *Journal of Econometrics* 47, 197–205.
- Ljungqvist, Lars and Thomas J. Sargent (2000) *Recursive Macroeconomic Theory*. Cambridge, Mass.: MIT Press.
- Lucas, Robert E. (1994) On the Welfare Cost of Inflation. Discussion Paper No. 394, Center for Economic Policy Research, Stanford University.
- Lucas, Robert E. and Nancy L. Stokey (1987) Money and interest rates in a cash-in-advance economy. *Econometrica* 55, 491–513.
- McGrattan, Ellen R. (1994) The Macroeconomic effects of distortionary taxation. *Journal of Monetary Economics* 33, 573–601.
- Mendoza, Enrique, Assaf Razin, and Linda Tesar (1994) Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics* 34, 297–323.
- Newey, Whitney K. and Kenneth D. West (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Sidrauski, Miguel (1967) Rational choice and patterns of growth in a monetary economy. *American Economic Review* 57, 534–544.
- Yaron, Amir (1996) Liquidity Shocks and International Asset Pricing. Manuscript, Carnegie Mellon University.

APPENDIX A

DATA

The OECD data is taken from the OECD publication *Main Economic Indicators*, various years.

A.1. UNITED STATES

The data on the money supply (M2) from 1947 to 1958 are from Friedman and Schwartz (1982). From 1959 onward, we use data from the Federal Reserve Bank of St. Louis. Because these sources contain monthly data, we take annual averages.

For the price level, we use the CPI series from the Federal Reserve Bank of St. Louis.

For nominal interest rates, we use the three-month treasury bill rate. Also, the monthly data are averaged in order to get annual data. Again, the source is the Federal Reserve Bank of St. Louis.

For government consumption, we use Citibase series GGE (nominal) and GGE82 (real). For private investment, we use the sum of GIF and GCD for the nominal series and the sum of GIF82 and GCD82 for the real series.

For private consumption, we use the sum of GCN and GCS and the sum of GCN82 and GCS82, respectively.

Income (whether nominal or real) is measured as the sum of private investment, government spending, and private consumption. This excludes increases in stocks and net exports and hence ensures that the identity $Y = C + I + G$ holds both in the data and the model.

Data on government debt are taken from Citibase, variable FBD. This series contains the gross debt of the federal government. Ideally, one would want data on the net debt (or assets) of the consolidated government, but such data are more difficult to construct.

Data on hours worked are taken from the Citibase variable LHOURS.

In order to get per capita values, we divide by the population as measured by the Citibase variable P16.

Data on capital and labor tax rates were given to us by Ellen McGrattan. The consumption tax rates were constructed by us, using the method proposed by Mendoza, Razin, and Tesar (1994).

A.2. SWEDEN

See Jonsson and Klein (1996) and Klein (1995) for details on the collection and calculation of the Swedish tax rates. The data on the money supply (M2) was provided by Lars Jonung. For nominal interest rates, we use the Central Bank Discount Rate, Statistical Yearbook, Central Bank of Sweden. All other variables are taken from the annual Swedish national accounts, Central Bureau of Statistics, Stockholm, various years.

APPENDIX B

TWO-STAGE GMM/SMM ESTIMATION

The following treatment is based on, but very slightly extends, the work of Lee and Ingram (1991) and Yaron (1996).

Suppose we have an empirical sequence $\{x_t\}_{t=1}^T$ and K independent simulated sequences $\{y_{t,i}(\hat{\beta})\}_{t=1}^N$, $i = 1, 2, \dots, K$. Define $n = \frac{NK}{T}$ and let n be constant whenever we let T and NK simultaneously tend to infinity. Suppose we find it convenient to estimate the parameters in two stages, where the first stage uses GMM and the second SMM. Formally,

define

$$\tilde{\beta} = \begin{bmatrix} \tilde{\beta}^1 \\ \tilde{\beta}^2 \end{bmatrix} \tag{21}$$

$$v_T^1(\tilde{\beta}^1) = \sum_{t=1}^T v^1(x_t, \tilde{\beta}^1) \tag{22}$$

$$m_T^1(\tilde{\beta}^1) = m^1(v_T^1(\tilde{\beta}^1)) \tag{23}$$

$$v_T^2 = \frac{1}{T} \sum_{t=1}^T v^2(x_t) \tag{24}$$

$$m_T^2 = m^2(v_T^2) \tag{25}$$

$$m_{j_1 \times 1}^1(\tilde{\beta}^1) = m^1(v_T(\tilde{\beta}^1)) \tag{26}$$

$$v_{N,i}^2(\tilde{\beta}) = \frac{1}{N} \sum_{t=1}^N v^2(y_{t,i}(\tilde{\beta})) \tag{27}$$

$$m_{N,i}^2(\tilde{\beta}) = m^2(v_{N,i}^2(\tilde{\beta})) \tag{28}$$

$$m_{j_2 \times 1}^2(\tilde{\beta}) = \frac{1}{K} \sum_{i=1}^K m_{N,i}^2(\tilde{\beta}). \tag{29}$$

Now suppose things are constructed so that (some model implies that), for some unique β ,

$$\text{plim}_{T \rightarrow \infty} [m_T^1(\beta^1)] = 0 \tag{30}$$

$$\text{plim}_{NK \rightarrow \infty} [m_{NK}^2(\beta)] - \text{plim}_{T \rightarrow \infty} (m_T^2) = 0. \tag{31}$$

This motivates the following estimator of β (the two-stage GMM/SMM estimator).

$$\beta_T^1 = \text{argmin}_{\beta^1} \{m_T^1(\tilde{\beta}^1)' W_T^1 m_T^1(\tilde{\beta}^1)\} \tag{32}$$

$$\beta_T^2 = \text{argmin}_{\beta^2} \{[m_{NK}^2(\beta_T^1, \tilde{\beta}^2) - m_T^2]' W_T^2 [m_{NK}^2(\beta_T^1, \tilde{\beta}^2) - m_T^2]\}, \tag{33}$$

where W_T^1 and W_T^2 are positive definite whose probability limits W^1 and W^2 are non-stochastic and positive definite. The subscript T indicates that these matrices may depend on the empirical observations. An alternative, which we ignore here for reasons of convenience, is to let the weighting matrix depend on simulations instead [see Yaron (1996)].

Subject to regularity conditions, this guarantees consistency. For asymptotic efficiency, we must choose W_T^1 and W_T^2 optimally. The optimal choice of W^1 is well described elsewhere [see, for example, Yaron (1996)].

The optimal choice of W^2 satisfies

$$\sqrt{T} [m_{NK}^2(\beta_T^1, \beta^2) - m_T^2] \xrightarrow{d} N[0, (W^2)^{-1}]. \tag{34}$$

Hence, our job is to find the asymptotic variance matrix of $(m_{NK}^2(\beta_T^1, \beta^2) - m_T^2)$. In order to do that, we begin by defining

$$v_T(\tilde{\beta}^1) = \begin{bmatrix} v_T^1(\tilde{\beta}^1) \\ v_T^2 \end{bmatrix},$$

(and similarly for other vectors of moments) and suppose that $\sqrt{T}[v_T(\beta^1)] \xrightarrow{d} N(v, S)$. Then we use our observations of $v(x_t, \beta_t^1)$ to estimate S by S_T . Now define Ω via

$$m_T \xrightarrow{d} N(m, \Omega). \tag{35}$$

Now, if m is smooth, we have a consistent estimator of Ω in

$$\Omega_T = \begin{bmatrix} \frac{\partial m^1(v_T^1)}{\partial (v_T^1)'} & 0 \\ 0 & \frac{\partial m^2(v_T^2)}{\partial (v_T^2)'} \end{bmatrix} S_T \begin{bmatrix} \frac{\partial m^1(v_T^1)}{\partial (v_T^1)'} & 0 \\ 0 & \frac{\partial m^2(v_T^2)}{\partial (v_T^2)'} \end{bmatrix}. \tag{36}$$

Finally, our desired variance matrix can be shown to be consistently estimated by

$$\begin{aligned} \Omega_T^* &= \begin{bmatrix} \Gamma_{21}(\Gamma'_{11} W_T \Gamma_{11})^{-1} \Gamma'_{11} W_T & I \\ W'_T \Gamma_{11}(\Gamma'_{11} W_T \Gamma_{11})^{-1} \Gamma'_{21} & I \end{bmatrix} \begin{bmatrix} (\Omega_T)_{11} & (\Omega_T)_{12} \\ (\Omega_T)'_{12} & (1 + \frac{1}{n})(\Omega_T)_{22} \end{bmatrix} \\ &\tag{37} \end{aligned}$$

Now, because Γ_{21} is a function of the second stage parameters, we get a circularity problem, which is solved in some standard way. See later for one option.

B.1. A SPECIAL CASE: THE FIRST STAGE IS EXACTLY IDENTIFIED

If $j_1 = k_1$ then Γ_{11} is square (but not necessarily symmetric) and if the first-stage moments really identify the first-stage parameters, then it is also invertible (with probability 1). Hence, our variance matrix reduces to

$$\Omega_T^* = \begin{bmatrix} \Gamma_{21} \Gamma_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} (\Omega_T)_{11} & (\Omega_T)_{12} \\ (\Omega_T)'_{12} & (1 + \frac{1}{n})(\Omega_T)_{22} \end{bmatrix} \begin{bmatrix} (\Gamma_{11}^{-1})' \Gamma'_{21} \\ I \end{bmatrix}. \tag{38}$$

B.1.1. A Special Case of the Special Case: The First Stage is a Regression

This is, of course, the case we consider in the main text. There we have an empirical sequence

$$\left\{ \begin{bmatrix} x_t^e \\ y_t^e \\ z_t^e \end{bmatrix} \right\}_{t=1}^T, \tag{39}$$

and K independent simulated sequences

$$\{z_{t,i}^s(\tilde{\beta})\}_{t=1}^N \quad i = 1, 2, \dots, K. \tag{40}$$

We will drop the superscripts e and s . Instead, subscripts T and N will represent empirical and simulated moments, respectively.

Now we partition

$$\tilde{\beta} = \begin{bmatrix} \tilde{\beta}^1 \\ \tilde{\beta}^2 \end{bmatrix}_{k_2 \times 1} = \begin{bmatrix} \mu \\ \text{vec}(\varphi) \\ \text{vech}(\Sigma) \\ \tilde{\beta}^2 \end{bmatrix}. \tag{41}$$

For convenience, define, for any sequence $\{w_t\}_{t=1}^T$,

$$\tilde{w}_t = w_t - \frac{1}{T} \sum_{t=1}^T w_t,$$

and set

$$\mu_T = \frac{1}{T} \sum_{t=1}^T y_t \tag{42}$$

$$\varphi_T = \left(\sum_{t=1}^T \tilde{y}_t \tilde{x}_t' \right) \left(\sum_{t=1}^T \tilde{x}_t \tilde{x}_t' \right)^{-1} \tag{43}$$

$$\Sigma_T = \frac{1}{T} \sum_{t=1}^T (y_t - \varphi_T x_t) (y_t - \varphi_T x_t)'. \tag{44}$$

Next we define

$$v_t = \begin{bmatrix} y_t - \mu_T \\ \text{vec}[\tilde{x}_t(\tilde{y}_t - \varphi_T x_t)'] \\ \text{vech}(y_t - \varphi_T x_t) (y_t - \varphi_T x_t)' - \text{vech}(\Sigma) \\ z_t \\ \text{vech}(\tilde{z}_t \tilde{z}_t') \end{bmatrix} \tag{45}$$

$$v_T = \frac{1}{T} \sum_{t=1}^T v_t, \tag{46}$$

and estimate, using the formula of Newey and West (1987), the asymptotic variance matrix of $\sqrt{T}v_T$. Call that matrix S and its estimator S_T . Then we partition according to

$$v_t^1 = \begin{bmatrix} y_t - \mu_T \\ \text{vec}[\tilde{x}_t(\tilde{y}_t - \varphi_T x_t)'] \\ \text{vech}(y_t - \varphi_T x_t) (y_t - \varphi_T x_t)' - \Sigma_T \end{bmatrix} \tag{47}$$

$$v_t^2 = \begin{bmatrix} z_t \\ \text{vech}(\tilde{z}_t \tilde{z}_t') \end{bmatrix}, \tag{48}$$

and define

$$v_T^1 = \frac{1}{T} \sum_{t=1}^T v_t^1 \tag{49}$$

$$v_T^2 = \frac{1}{T} \sum_{t=1}^T v_t^2. \tag{50}$$

Next we define

$$\Gamma_{11} = \begin{bmatrix} -I_m & & \\ & -I_m \otimes \left(\frac{1}{T} \sum_{t=1}^T \tilde{x}_t \tilde{x}_t' \right) & \\ & & -I_{m(m+1)/2} \end{bmatrix}. \tag{51}$$

(Note that Γ_{11} is symmetric in this case.) Then we define

$$m_T^2 = m^2(v_T^2) \tag{52}$$

$$\Omega_T = \begin{bmatrix} I_{m(r+1)+m(m+1)/2} & 0 \\ 0 & \frac{\partial m^2(v_T^2)}{\partial (v^2)'} \end{bmatrix} S_T \begin{bmatrix} I_{m(r+1)+m(m+1)/2} & 0 \\ 0 & \frac{\partial m^2(v_T^2)}{\partial (v^2)} \end{bmatrix} \tag{53}$$

$$m_N^2(\tilde{\beta}) = \frac{1}{K} \sum_{j=1}^K m_{N,i}^2(\tilde{\beta}), \tag{54}$$

($m_{N,i}^2$ being defined in the obvious way) and

$$\beta_{TN}^2 = \operatorname{argmin}_{\beta^2} \{ [m_N^2(\beta_T^1, \beta^2) - m_T^2]' [m_N^2(\beta_T^1, \beta^2) - m_T^2] \}. \tag{55}$$

Next we define (and calculate using numerical derivatives)

$$\Gamma_{21} = \frac{\partial m_N^2(\beta_T^1, \beta_{TN}^2)}{\partial (\beta^1)'} \tag{56}$$

$$\Omega_T^* = \begin{bmatrix} \Gamma_{21} \Gamma_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} (\Omega_T)_{11} & (\Omega_T)_{12} \\ (\Omega_T)'_{12} & (1 + \frac{1}{n})(\Omega_T)_{22} \end{bmatrix} \begin{bmatrix} \Gamma_{11}^{-1} \Gamma_{21}' \\ I \end{bmatrix} \tag{57}$$

$$J_{TN}(\tilde{\beta}^1, \tilde{\beta}^2) = [m_N^2(\tilde{\beta}^1, \tilde{\beta}^2) - m_T^2]' (\Omega_T^*)^{-1} [m_N^2(\tilde{\beta}^1, \tilde{\beta}^2) - m_T^2] \tag{58}$$

$$(\beta_{TN}^2)^* = \operatorname{argmin}_{\beta^2} [J_{TN}(\beta_T^1, \tilde{\beta}^2)],$$

and, finally

$$T \cdot J_{TN}(\beta_T^1, \beta_{TN}^2) \xrightarrow{d} \chi_{(j_2-k_2)}^2. \tag{59}$$

As noted earlier, we need estimates of the second stage parameters in order to find the derivatives that are used to construct the optimal weighting matrix. To find consistent estimates of β^2 for this purpose, we use the identity matrix as the weighting matrix rather than the optimal one. Given these parameter estimates, we then calculate the optimal weighting matrix and estimate β^2 once more.