Centre interuniversitaire de recherche
en économie quantitative

CIREQ

Cahier 15-2014

*Indirect Inference Estimation of
Nonlinear Dynamic General Equilibrium Models :
With an Application to Asset Pricing under Skewness Risk*

Francisco RUGE-MURCIA

Cahier 15-2014

Indirect Inference Estimation of Nonlinear Dynamic General Equilibrium Models : With an Application to Asset Pricing under Skewness Risk

Francisco RUGE-MURCIA

Dépôt légal - Bibliothèque nationale du Canada, 2014, ISSN 0821-4441
Dépôt légal - Bibliothèque et Archives nationales du Québec, 2014
ISBN-13 : 978-2-89382-675-2

Indirect Inference Estimation of Nonlinear Dynamic General Equilibrium Models: With an Application to Asset Pricing under Skewness Risk

Francisco Ruge-Murcia*

December 2014

Abstract

This paper proposes a nonlinear impulse-response matching procedure explicitly designed to estimate nonlinear dynamic models, and illustrates its applicability by estimating a macro-finance model of asset pricing under skewness risk. As auxiliary model, a new class of nonlinear vector autoregressions (NVAR) based on Mittnik (1990) is proposed.

JEL classification: C51, C58
Keywords: nonlinear vector autoregression, nonlinear impulse responses, skewness risk.

*This research received the financial support of the Social Sciences and Humanities Research Council of Canada.
Correspondence: Department of Economics, McGill University, 855 Sherbrooke Street West, Montréal, Québec, Canada, H3A 2T7. E-mail: francisco.ruge-murcia@mcgill.ca

1. Introduction

Matching impulse responses is a widely used indirect inference procedure to estimate dynamic general equilibrium models. Theoretical foundations for this estimation strategy are provided in the seminal contributions by Gourieroux, Monfort, and Renault (1993) and Smith (1993). Previous applications include, among others, Rotemberg and Woodford (1998), Christiano, Eichenbaum, and Evans (2005), Iacoviello (2005), Boivin and Giannoni (2006), Uribe and Yue (2006), DiCecio and Nelson (2007), Dupor, Han, and Tsai (2007), Jorda and Kozicki (2007), and Altig, Christiano, Eichenbaum, and Linde (2011). Information criteria for the selection of impulse responses are developed by Hall, Inoue, Nason, and Rossi (2012). Although indirect inference was proposed initially as a method to estimate nonlinear models (see the examples in Gourieroux, Monfort, and Renault, 1993), most of the above applications concern linear or linearized models where impulse responses are independent of the sign, size, and timing of the shock. Since the response to a shock of size $+1$ (say, standard deviation) is one-half the response to a shock of size $+2$, the mirror image of a response to a shock of size -1, and independent of the moment the shock takes place, it is sufficient to consider only one shock to describe the model dynamics.

With advances in nonlinear solution methods and the increase in computing power, nonlinear dynamic models in macroeconomics and finance are now often estimated, rather than calibrated. However, the use of impulse-response matching in this setup must address the fact that in nonlinear systems a single response does not completely characterize the dynamic effects of a shock. Instead, the effect depends on the sign, size, and timing of the shock (see Gallant, Rossi, and Tauchen, 1993, and Koop, Pesaran, and Potter, 1996). Of course, under the conditions in Smith (1993) (see, also, Dridi, Guay, and Renault, 2007) ignoring nonlinearity—that is, using a linear vector autoregression (VAR) as auxiliary model and a single impulse response as binding function—delivers consistent estimates of the structural parameters. The point made here is simply that this approach is inefficient when the data generating process (DGP) is truly nonlinear.

As an alternative, I propose a novel class of nonlinear vector autoregressions (NVAR) based on Mittnik (1990) that can serve as auxiliary model and, like the economic model, generates nonlinear impulse responses. Nonlinear impulse responses exploit information on the curvature of the model to provide more comprehensive information about the model dynamics than a single linear response does. Since the efficiency of the indirect inference estimator increases as the auxiliary model approximates better the actual DGP (Smith, 2008), using the proposed NVAR rather than a VAR should deliver gains in statistical efficiency. I provide Monte-Carlo evidence that this is indeed the case. In addition, I argue that some of the identification issues raised in Canova and Sala

(2009)—for example, under-identification due to the normalization of the shock used to compute responses in linear models and the fact that information about the steady state is not used in the estimation—are less severe in nonlinear than in linear models.

As an illustration of the proposed method, this paper estimates a nonlinear macro-finance model of asset pricing under skewness risk. This application is important in its own right. Traders have Epstein-Zin preferences (Epstein and Zin, 1989) and live in a production economy where firms use labor and capital as inputs. The financial assets are shares or claims on the dividends of firms, and the source of risk is a productivity shock. Compared with earlier research that typically treats this shock as a latent variable, this project uses the series of total factor productivity for the U.S. constructed by Fernald (2014) as one of the observable variables. I provide evidence of significant departures from Gaussianity and, in particular, show that productivity innovations are negatively skewed and leptokurtic. For this reason, the model is estimated under the assumption that innovations are drawn from an asymmetric Gumbel distribution. Since shareholders are subject to potentially large negative realizations from the long tail of the distribution and such realization reduce the return on capital, they are subject to skewness risk. The model is solved using a nonlinear perturbation method that makes explicit the dependence of consumption and asset returns on the second- and third-order moments of the productivity innovations.

Although the quantity of aggregate risk in the canonical growth model is too small to explain the equity premium puzzle (a point previously made by Campanale, Castro and Clementi, 2010), it shown that the price of skewness risk is one order of magnitude larger than the price of variance risk, that the nonlinear model can endogenously generate ARCH effects in asset returns, and that the responses of consumption and asset returns to productivity shocks are asymmetric, with negative shocks inducing larger responses than positive shocks.

The paper is organized as follows. Section 2 develops a macro-finance model of asset pricing where consumption and asset prices are both endogenous and driven by asymmetric productivity shocks. Section 3 proposes a nonlinear time series model specifically designed to play the role of auxiliary model in the indirect inference estimation of nonlinear dynamic general equilibrium models, and discusses the use of nonlinear impulse responses as binding functions. Section 4 describes the data used to estimate the model, reports evidence of significant departures from Gaussianity, reports parameter estimates, and examines the economic implications of asymmetric shocks and skewness risk for consumption and stock returns. Section 5 concludes.

2. Asset Prices in a Production Economy with Skewness Risk

The representative trader has recursive preferences over consumption (Epstein and Zin, 1989),

$$U_t = \left((1-\beta) (C_t)^{1-1/\psi} + \beta \left(E_t \left(U_{t+1}^{1-\gamma} \right) \right)^{(1-1/\psi)/(1-\gamma)} \right)^{1/(1-1/\psi))}, \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, C_t is consumption, E_t is the expectation conditional on information available at time t, γ is the coefficient of risk aversion, and ψ is the intertemporal elasticity of substitution (IES). Time is discrete. In every period, the trader supplies a fixed time endowment in a competitive labor market and participates in a financial market where shares can be bought and sold. The trader's budget constraint is

$$C_t + Q_t S_{t+1} = X_t H + (Q_t + D_t) S_t, \tag{2}$$

where Q_t is the price of a share, S_t is the number of shares, X_t is the hourly wage, H is hours worked, and D_t is dividends. Since consumption is the numeraire, Q_t and X_t are real prices in terms of units of consumption. The Euler equation that characterizes the trader's utility maximization is

$$\beta E_t \left(\Lambda_{t,t+1} R_{t+1} \right) = 1, \tag{3}$$

where $\Lambda_{t,t+1} = (V_{t+1}/W_t)^{1/\psi - \gamma} (C_{t+1}/C_t)^{-1/\psi}$ is the pricing kernel, $V_t \equiv \max U_t$ is the value function, $W_t \equiv E_t V_{t+1}$ is the certainty-equivalent future utility, and

$$R_{t+1} = (Q_{t+1} + D_{t+1})/Q_t \tag{4}$$

is the gross return on a share bought at time t.

Euler equations of this form have been extensively studied in the asset-pricing literature. One approach involves estimating the parameters using, for example, the generalized method of moments and testing the over-identifying restrictions of the model. Well-known examples are Hansen and Singleton (1982) for expected utility, and Epstein and Zin (1991) for recursive utility. Another approach involves assuming that the arguments inside the expectations operator are jointly Log-normal and conditionally homoskedastic to obtain a linear pricing function with a risk-adjustment factor, as in Jerman (1998). This pricing function is tractable but it rules out time-variation in the risk premia (because the risk-adjustment factor is constant) and assumes away the contribution of higher-order moments (because the factor is proportional to the variance only). Rather than focusing on the Euler equation alone, the strategy in this paper is to solve the complete model and characterize its dynamics using a nonlinear approximation to the policy functions. As we will see below, policy functions are attractive because they make explicit the dependence of prices and quantities on the state variables of the model and on the moments of the innovations. In turn, this will be useful in informing the restrictions that I will impose on the auxiliary model for the indirect inference estimation of the model.

The representative firm produces output using the technology

$$Y_t = Z_t \left(H_t\right)^{1-\alpha} \left(K_t\right)^{\alpha}, \tag{5}$$

where Y_t is output, Z_t is a productivity shock, H_t is labor input, K_t is capital, and $\alpha \in (0,1)$ is a parameter. The shock follows the process

$$\ln Z_t = \rho(L) \ln Z_{t-1} + \epsilon_t, \tag{6}$$

where L is the lag operator, $\rho(L) = (1 - \rho_1 L - \ldots - \rho_v L^v)$ is a polynomial with roots outside the unit circle, and ϵ_t is an independent and identically distributed (i.i.d.) innovation with mean zero, constant conditional variance, and non-zero skewness. This specification relaxes two standard assumptions in earlier literature. First, the time series process of the productivity shock is generalized from an AR(1) to an AR(v). This is important because the latter process allows many possible shapes for impulse responses while the former only permits geometric decay. In particular, the AR(v) process can generate hump-shaped impulse responses while the AR(1) cannot. Below, I show that this is important to capture the dynamic effects of productivity shocks on consumption and stock returns.

Second, and more importantly, by allowing non-zero skewness, this specification relaxes the assumption that shocks are symmetric. This assumption is implicit in models with t or Normal shocks and implies that positive and negative realizations of the same magnitude are equally likely. Instead, in an economy with asymmetric shocks traders are subject to skewness risk—i.e., the possibility of large realizations from the long tail of the shock distribution. In the empirical part of this project, I assume that productivity innovations are drawn from a Gumbel distribution (Gumbel, 1935). This distribution is used to model the minima (or the maxima) of a sample of i.i.d. realizations and is the special case of the Generalized Extreme Value (GEV) distribution where the shape parameter is zero.[1] Figure 1 plots the probability distribution function of the Gumbel distribution (thick line) and compares it with the one of a Normal distribution with the same variance (thin line). In this figure, the units in the horizontal axis are standard deviations from the mean. Notice that the Gumbel distribution is negatively skewed, with more probability mass in the left, and less probability mass in the right, than the Normal distribution. This means

[1] In preliminary work I estimated the parameters of the GEV distribution of productivity innovations and found the shape parameter to be quantitatively and statistically close to zero. For this reason, I focus on the Gumbel distribution in section 4.

that large negative productivity innovations are more likely than a large positive innovations of the same magnitude. Section 4.1 below provides statistical evidence of significant departures from Gaussianity and negative skewness in U.S. productivity innovations.

The firm owns directly its capital stock. The law of motion for capital is

$$K_{t+1} = (1 - \delta)K_t + I_t, \tag{7}$$

where $\delta \in (0, 1)$ is the rate of depreciation and I_t is investment. Adjusting the capital stock involves a convex cost that takes the form

$$\Phi_t = \Phi(I_t/K_t) = (\phi/2)(I_t/K_t - \delta)^2 K_t, \tag{8}$$

where $\phi \geq 0$ is a cost parameter.[2] The representative firm optimally chooses inputs to maximize

$$E_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau} \Lambda_{\tau,t} D_t, \tag{9}$$

subject to the technology (5) and the law of motion (7). Profits, which are transferred to shareholders in the form of dividends, are given by

$$D_t = Y_t - X_t H_t - I_t - \Phi_t. \tag{10}$$

Notice that due to the assumption of constant returns to scale, dividends are the return on capital net of new investment and adjustment costs.

The equilibrium is an allocation for the trader $\mathcal{C} = \{C_t, S_{t+1}\}_{t=\tau}^{\infty}$, an allocation for the firm $\mathcal{Y} = \{Y_t, H_t, K_{t+1}\}_{t=\tau}^{\infty}$ and a price system $\mathcal{P} = \{Q_t, W_t\}_{t=\tau}^{\infty}$ such that given the price system: (i) the allocation $\mathcal{C}$ solves the trader's problem; (ii) the allocation $\mathcal{Y}$ solves the firms's problem; (iii) share holdings add up to 1: $S_t = 1$; (iv) the labor market clears: $H_t = H$; and (v) the goods market clears: $C_t + I_t + \Phi_t = Y_t$.

This model does not have an exact analytical solution but it is straightforward to apply numerical methods to compute an approximate nonlinear solution. In this paper, I use a perturbation method based on Jin and Judd (2002) that consists in taking a third-order polynomial expansion of the policy functions around the deterministic steady state and characterizing the local dynamics. Caldera, Fernández-Villaverde, and Rubio-Ramírez (2012) show that for models with recursive preferences, a third-order perturbation is as accurate as projection methods in the range of interest

[2] In preliminary work I estimated the model using the Linex function due to Varian (1974). This function allows costs to depend on both the magnitude and sign of the investment rate and nests the quadratic function (8) in the special case where the asymmetry parameter tends to zero. The indirect inference estimator of the asymmetry parameter was positive, which suggests some irreversibility in investment, but it was quantitatively small and, thus, results are basically the same as those reported here for the quadratic function.

but is much faster computationally. I implement the solution method using the codes described in Ruge-Murcia (2012).

In general terms, a policy function takes the form $f(s_t, \sigma)$ where s_t is a vector of state variables and σ is a perturbation parameter. In this model, the state variables are the capital stock and current and lagged realizations of the productivity shock. That is, $s_t = [K_t\ Z_t\ Z_{t-1}\ \ldots\ Z_{t-(v-1)}]'$. The goal is to approximate $f(s_t, \sigma)$ using a third-order polynomial expansion around the deterministic steady state where $s_t = s$ and $\sigma = 0$. This approximation can be written using tensor notation as

$$\begin{aligned}[f(s_t,\sigma)]^j \quad = \quad & [f(s,0)]^j + [f_s(s,0)]^j_a[(s_t-s)]^a \\ & +(1/2)[f_{ss}(s,0)]^j_{ab}[(s_t-s)]^a[(s_t-s)]^b \\ & +(1/2)[f_{\sigma\sigma}(s,0)]^j[\sigma][\sigma] \\ & +(1/6)[f_{sss}(s,0)]^j_{abc}[(s_t-s)]^a[(s_t-s)]^b[(s_t-s)]^c \\ & +(1/2)[f_{s\sigma\sigma}(s,0)]^j_a[(s_t-s)]^a[\sigma][\sigma] \\ & +(1/6)[f_{\sigma\sigma\sigma}(s,0)]^j[\sigma][\sigma][\sigma],\end{aligned} \tag{11}$$

where j refers to the j-th variables in the model (say, stock returns), a, b, and c are indices, and terms like $[f_s(s,0)]^j_a$ are coefficients that depend on the structural parameters of the model.[3]

Since the model solution motivates the restrictions that I impose on the auxiliary model, it is useful to examine its structure in some detail. First, the policy function includes linear, quadratic, and cubic terms in the state variables. Second, the policy function depends on higher-order moments of the shock innovations in the form of a risk-adjustment factor that is proportional to both the variance and skewness. The effect of the skewness can be seen in the term $(1/6)[f_{\sigma\sigma\sigma}(s,0)]^j[\sigma][\sigma][\sigma]$, which is non-zero in the case where innovations follow an asymmetric distribution because skewness is non-zero in this case.[4] In addition to this direct effect on the ergodic mean of the variables, the skewness of the innovations also effects the dynamics through the asymmetry it induces in the state (and, hence, in the control) variables. All these observations imply that a reduced-form representation of the policy function is a third-order polynomial in the state variables with a non-zero intercept.

[3]In writing (11), I have used the intermediate results $[f_{s\sigma}(s,0)]^j_a = [f_{\sigma s}(s,0)]^j_a = 0$ (Schmitt-Grohé and Uribe, 2004, p. 763), $[f_{ss\sigma}(s,0)]^j_{ab} = [f_{\sigma ss}(s,0)]^j_{ab} = [f_{s\sigma s}(s,0)]^j_{ab} = 0$ (Ruge-Murcia, 2012, p. 936), and $[f_{\sigma\sigma s}(s,0)]^j_a = [f_{\sigma s\sigma}(s,0)]^j_a = [f_{s\sigma\sigma}(s,0)]^j_a$ (Clairaut's theorem).

[4]Note that, strictly speaking, $[\sigma][\sigma][\sigma]$ is the third moment of the innovations, rather than skewness *per-se*, which is this third moment divided by the cube of the standard deviation.

3. Nonlinear Impulse-Response Matching

Indirect inference requires the selection of an auxiliary model and binding function, and these choices typically depend on the economic model to be estimated. In this section, I propose a new nonlinear time series model based on Mittnik (1990) that generalizes linear vector autoregressions and is well-suited to play the role of auxiliary model in the indirect inference estimation of nonlinear dynamic general equilibrium models. I refer to this model as a nonlinear vector autoregression and write it as NVAR(n, p) where n is the number of lags in the autoregression and p is the polynomial order. I also discuss the use of nonlinear impulse responses as binding function and describe the optimization problem that delivers an indirect inference estimator in this setup.

3.1 Auxiliary Model

A general NVAR is written as

$$y_t = \Omega_0 + \Omega_1 x_t + \Omega_2(x_t \otimes x_t) + \ldots + \Omega_p(x_t \otimes \ldots \otimes x_t) + \epsilon_t, \tag{12}$$

where y_t is a $k \times 1$ vector of observable variables, x_t is a $nk \times 1$ vector with n lags of each of the k variables in y_t, ϵ_t is a $k \times 1$ vector of residuals, Ω_0 is a $k \times 1$ vector of constants, Ω_i for $i = 1, 2, \ldots, p$ are conformable matrices with fixed parameters, and $\otimes$ denotes the Kronecker tensor product. This multivariate NVAR(n, p) model is closely related to the univariate Generalized Autoregression (GAR) due to Mittnik (1990) where the conditional mean of the variable is a function of its lagged values and a polynomial in lagged cross-products over time. Mittnik notes that since the relation between the variable and its lags is linear and the error term is separable, it is possible to use least-squared methods to estimate the GAR coefficients. These conditions also hold for the NVAR and in the empirical application and Monte-Carlo experiments below, I use ordinary least squares (OLS) equation by equation to estimate the NVAR coefficients. Thus, the NVAR meets one of the key attributes that an auxiliary model should have for the purpose of indirect inference, namely that it should be relatively easy to estimate.

Using a completely unrestricted formulation of (12) for the estimation of the model is impractical for several reasons. First, the number of parameters increases rapidly with the number of variables, the number of lags, and polynomial order of the NVAR. Second, it is difficult to impose conditions to insure the stability and stationarity of a general NVAR. Finally, an unrestricted NVAR ignores features of the economic model that can provide a tighter link between structural and auxiliary models, and rule out explosive paths. For these reasons, it is desirable to impose some restrictions on the elements of the matrices Ω_i for $i = 1, 2, \ldots, p$. In what follows, I illustrate some of these restrictions in the context of the asset pricing model in section 2 and assuming that the econometrician has access to data on productivity, consumption, and stock returns to estimate the model. That is, $y_t = [Z_t, C_t, R_t]'$.

Some restrictions on the NVAR are motivated by the process of the productivity shock in the economic model (see (6)). Since Z_t follows an exogenous, linear process that depends only on v of its lags, one could restrict the coefficients of lagged higher-order terms in productivity and of all terms in consumption and returns in the first equation of (12) to be zero. In addition to this model-based argument, there is also an statistical one. During the estimation procedure, artificial data are generated from the structural model to compute synthetic NVARs and impulse-responses. Since (6) holds in the model, the coefficients of the additional terms in the auxiliary model should not be statistically different from zero. Imposing these restrictions on (12) delivers sharper estimates of the NVAR coefficients and impulse-responses, and avoids the extra "noise" that would arise because the coefficients of the superfluous variables will not be identically equal to zero in a finite sample. With these restrictions imposed, the first equation in the NVAR corresponds exactly to (6) and is stationary and ergodic by assumption.

Other set of restrictions are motivated by the model solution. The third-order perturbation solution (11) shows that consumption and returns are functions of a third-order polynomial on the state variables ($Z_t, Z_{t-1}, \ldots, Z_{t-(v-1)}$, and K_t). Thus, C_t and R_t in the NVAR should be functions of (at least) linear, quadratic and cubic terms in current and $v-1$ lags of productivity. Since under the model solution, choice variables depend solely on state variables, the coefficients of all terms in consumption and returns in the last two equations of the NVAR should be set to zero. Because the process for productivity is stationary and all other variables are specified as functions of productivity only, the above restrictions imply that the NVAR computed using data from a stationary model has a unique steady state and no explosive paths regardless of the shock size.

In related work, Aruoba, Bocola, and Schorfheide (2014) map the second-order perturbation solution for dynamic general equilibrium models into a quadratic autoregression (QAR) and characterize its impulse-responses. (The second-order perturbation solution is the special case of (11) without cubic terms.) My research complements their work by proposing a general nonlinear model that accommodates perturbation solutions of any order and can be used in a multivariate environment. Barnichon and Matthes (2014) construct nonlinear impulse responses by using Gaussian basis functions to parameterize the coefficients of an atheoretical moving average representation of a system. Depending on the function parameters, impulse responses may be asymmetric, hump-shaped, and/or display overshooting and oscillations. Similarly, the impulse-responses of a NVAR can display these features depending on the lag length and the order of the polynomial.

3.2 Binding Function

The binding function maps the parameters of the economic model into those of the auxiliary model. This role is played here by the nonlinear impulse responses. Recall that in linear models, impulse responses are exactly proportional to the sign and size of the shock, and independent of its timing. Thus, one impulse response (of any size or sign) is sufficient to describe the model dynamics. However, in this paper the auxiliary model is a nonlinear vector autoregression whose impulse responses depend on the sign, size, and timing of the shock in a non-trivial manner. Fortunately, methods to compute impulses responses in nonlinear systems have been proposed in earlier literature by Gallant, Rossi, and Tauchen (1993), Koop, Pesaran, and Potter (1996), and Pesaran and Shin (1998). As an alternative one may also consider computing impulse responses to shocks located in different percentiles of the innovation distribution, which are of different sign and size by construction. Any of these methods would provide a suitable characterization of the model responses for the purpose of indirect inference estimation.

3.3 Statistical Problem

With the above elements in place, we are ready to formulate the indirect inference estimator in this setup. Consider a nonlinear model with unknown parameters $\theta \in \Theta$, where θ is a $q \times 1$ vector and $\Theta \subset \Re^q$ is a compact and bounded set. The econometrician has at her disposal a sample of $T + n$ observations of N data series to estimate the model. Denote this sample by $\{w_t\}_{t=-(n-1)}^{T}$ where w_t a $N \times 1$ vector and assume that w_t is stationary and ergodic, possibly as a result of a prior transformation of the raw data by means of a detrending procedure. Denote by $y_t(\theta)$ the $N \times 1$ vector with artificial data simulated from the economic model using parameter values θ and assume that $y_t(\theta)$ is stationary and ergodic for all $\theta \in \Theta$. The size of the simulated sample is $\lambda T + n - 1$ with $\lambda \geq 1$ to stress the fact that, in general, the simulated sample is larger than the size of the actual sample. Under the null hypothesis, there exists a unique $\theta_0 \in \Theta$, where θ_0 is an interior point of Θ, such that the random sequences $\{w_t\}_{t=-(n-1)}^{T}$ and $\{y_t(\theta)\}_{t=-(n-1)}^{\lambda T}$ have identical stationary distributions.

Using the simulated sample $\{y_t(\theta)\}_{t=-(n-1)}^{\lambda T}$, construct the nonlinear vector autoregression (12) with appropriate restrictions to insure its stationarity and ergodicity. Denote the estimates of the NVAR(n, p) coefficients as $\hat{\varrho}(\theta, \lambda T)$ and the impulse responses generated by this system as $\hat{\gamma}(\theta, \lambda T, h)$, where h is the horizon of the responses. In terms of the notation in section 3.1, $\hat{\varrho}(\theta, \lambda T)$ contains the elements of $\hat{\Omega}_0(\theta, \lambda T)$, $\hat{\Omega}_1(\theta, \lambda T)$, $\hat{\Omega}_2(\theta, \lambda T), \ldots,$ and $\hat{\Omega}_p(\theta, \lambda T)$. Note that the responses in $\hat{\gamma}(\theta, \lambda T, h)$ are rearranged as a $\mu h \times 1$ vector where μ is the number of responses computed. Similarly for the actual data, construct the same nonlinear vector autoregression and compute the estimates $\hat{\varrho}(T)$ and the impulse responses $\hat{\gamma}(T,h)$. Under the regularity conditions in Smith (1993, sec. 2.3), $\hat{\varrho}(\theta,\lambda T)$ and $\hat{\gamma}(\theta,\lambda T,h)$ converge in probability to their pseudo-true values $\varrho(\theta,\lambda T)$ and $\gamma(\theta,\lambda T,h)$ as $\lambda T\to\infty$, and $\hat{\varrho}(T)$ and $\hat{\gamma}(T,h)$ converge in probability to their pseudo-true values ϱ_0 and $\gamma_0(h)$ as $T\to\infty$. Then, the indirect inference estimator is

$$\hat{\theta} = \arg\min \left(\hat{\gamma}(T,h) - \hat{\gamma}(\theta,\lambda T,h)\right)' W \left(\hat{\gamma}(T,h) - \hat{\gamma}(\theta,\lambda T,h)\right), \tag{13}$$

where W is a $\mu h\times\mu h$ is a weighting matrix. Intuitively, the indirect inference estimator minimizes the weighted distance between the impulse responses predicted by the auxiliary NVAR model computed using actual and artificial data simulated from the economic model.[5] Under assumptions corresponding to 1 to 7, 9, and 10 in Smith (1993),

$$\sqrt{T}(\widehat{\theta}-\theta_0)\to N(0,(1+1/\lambda)(J'WJ)^{-1}J'WSWJ(J'WJ)^{-1}), \tag{14}$$

where J is the $\mu h\times q$ Jacobian matrix, which is assumed to be finite and of full column rank, and S is the $\mu h\times\mu h$ long-run variance-covariance matrix. The multiplicative term $(1+1/\lambda)>1$ captures the effect of simulation uncertainty but its magnitude decreases geometrically with λ. In the special case where the weighting matrix $W=S^{-1}$, the asymptotic distribution of the indirect inference estimator becomes

$$\sqrt{T}(\widehat{\theta}-\theta_0)\to N(0,(1+1/\lambda)(J'S^{-1}J)^{-1}).$$

3.4 Monte-Carlo Experiments

In order to examine the performance of the proposed indirect inference estimator, I carry out a limited number of Monte-Carlo experiments. The data generating process are versions of the model in section 2 with different degrees of nonlinearity. Since the solution and estimation of the model can be time-consuming,[6] I focus on a small number of parameters and in specifications that can be accurately solved using a second-order perturbation. In particular, I consider cases where traders have preferences with constant relative risk aversion (that is, $\gamma=1/\psi$) and productivity follows an AR(1) process with normally distributed innovations.

[5] An alternative indirect inference estimator that is closely related to the extended method of simulated moments (EMSM) proposed by Smith (1993) is

$$\hat{\theta} = \arg\min(\hat{\varrho}(T)-\hat{\varrho}(\theta,\lambda T))'V(\hat{\varrho}(T)-\hat{\varrho}(\theta,\lambda T)),$$

where V is a weighting matrix. In order to keep the scope of this project manageable, I abstain from exploring further the properties of this estimator.

[6] To see this, notice that in order to compute gradients, a model with q parameters needs to be solved and simulated at least q times in each iteration of the algorithm that optimizes the statistical objective function.

The estimated parameters are the autoregressive coefficients of the productivity shock (ρ), the standard deviation of the productivity innovations (η), the coefficient of risk aversion (γ), and the capital adjustment cost (ϕ). In all experiments, the discount factor (β) is fixed to 0.9988, the depreciation rate is fixed to $\delta = 0.0075$, and the labor share $(1 - \alpha)$ is fixed to 0.66. Since the period in the model is taken to be one month, the value $\beta = 0.9988$ implies a real interest rate of about 1.45% at the annual rate, which is close to the average real return on Treasury bills during the sample period (see below). A monthly depreciation rate of 0.0075 implies an annual rate of about 9% and the value $(1 - \alpha) = 0.66$ is consistent with data from the National Income and Product Accounts (NIPA) that show that the share of labor in total income is approximately this value.

All experiments are based on 100 replications using the identity matrix as weighting matrix and a horizon of 50 periods for the impulse responses. I study samples with 200 and 500 observations, and artificial samples that are 10 times larger than actual one (i.e., $\lambda = 10$). The simulations of the nonlinear model, both for data generation and for indirect inference estimation, are based on the pruned version of the model as suggested by Kim, Kim, Schaumburg, and Sims (2008). Pruning has the attractive feature that it insures that the second-order perturbation has stable higher-order dynamics if the linear dynamics are stable. The data series are productivity, consumption and stock returns in deviation from their deterministic steady-state values.

Results are reported in tables 1 and 2. In the tables, True is the parameter value used to generate the data, Mean and Median are, respectively, the mean and the median of the estimated parameter values, and s.e. is the mean asymptotic standard error. Linear refers to the model estimated using a VAR(1) as auxiliary model and generator of impulse responses, while Nonlinear refers to the model estimated using a NVAR(1,2). In both cases, I impose the theory-based restrictions discussed in section 3.1.

Results in these tables support two conclusions. First, regardless of whether one uses the linear or the nonlinear vector autoregression as auxiliary model, indirect inference delivers point estimates that are quantitatively close to the true parameter values, even when the sample size is relatively small. This result is (of course) due to the fact that, in both cases, the indirect inference estimator is consistent for the structural parameters. Second, and more important from the perspective of this project, standard errors are considerably smaller in experiments where the auxiliary model is the NVAR rather than the VAR. This result is an illustration of the result (see, e.g., Smith, 2008) that the efficiency of the indirect inference estimator increases as the auxiliary model approximates better the data generating process (DGP), reaching the asymptotic efficiency of maximum likelihood in the limit. In this case, the true economic model is nonlinear and, hence, the appropriately specified

NVAR is a better approximation to the DGP than the linear VAR. This accounts for the efficiency gains reported in tables 1 and 2.

Some of the identification issues raised in Canova and Sala (2009) are likely to be less severe in nonlinear models, as well. Referring specifically to impulse-response matching, Canova and Sala note possible under-identification because the shock used to compute impulse responses in linear models is normalized and because this procedure does not exploit information about the steady state. Instead, in the case of nonlinear models, shocks are not normalized and using shocks of multiple size and sign allows the researcher to exploit the relation between model curvature and impulse responses to better identify the structural parameters. In addition, since the mean deviation from the deterministic steady state is not zero in nonlinear models (while it is zero in linear models, by construction), the researcher can use information on the first-moments of the data, for example, by specifying that shocks take place when the system is at the stochastic steady state, where all variables are at the mean of their ergodic distributions.

4. Application

This section illustrates the application of nonlinear impulse-response matching to the estimation of the asset pricing model in section 2. In particular, the section provides motivating evidence for relaxing the restrictive assumption of normally distributed productivity shocks, discusses the implementation of nonlinear impulse-response matching using the novel nonlinear vector autoregression, reports parameter estimates, and examines the economic implications of the model.

4.1 Data

The asset pricing model is estimated using monthly observations of productivity, real per-capita consumption, and stock returns between January 1971 and June 2014. The sample starts in January 1971 because this is the first available observation of the market index used to compute stock returns, and ends in June 2014 because this was the latest available observation at the time data was collected. The number of observations in the sample is 522. Productivity is measured using the total factor productivity series constructed by John Fernald (Fernald, 2014) and available from the Web site of the Federal Reserve Bank of San Francisco (*www.frbsf.org/economic-research/total-factor-productivity-tfp*). The raw data is a growth rate at the quarterly frequency, but I time aggregate the data to obtain a productivity index in levels and then use cubic interpolation to construct a monthly series. Consumption is measured by personal consumption expenditures on nondurable goods and services divided by the consumer price index (CPI) and the civilian non-institutional population. The latter is defined as persons 16 years of age who are neither inmates of institutions nor on active duty in the Armed Forces.[7] Consumption and productivity are quadratically detrended. Stock returns are measured by the monthly real return on the Wilshire 5000 Total Market Index, which include reinvested dividends. The source for all data (other than the productivity series) is the Federal Reserve Bank of St. Louis (*www.stlouisfed.org*).

4.2 Non-Normality

This section reports evidence of statistically significant departures from normality in productivity, consumption, and stock returns. The observation that stock returns are non-Gaussian is well known in the literature (see, among many others, Mandelbrot and Hudson, 2004), but I am not aware of research documenting the non-Gaussianity of productivity shocks and its innovations. This is important for the large literature in macroeconomics that uses productivity shocks as a key model ingredient and typically assumes that they are drawn from a Normal distribution. I focus initially on the series and sample used to estimate the model and then, in order to show that conclusions are robust to the detrending procedure and sample size, I report additional results based on growth rates and a longer sample size.

Figure 2 plots the histograms of productivity innovations, consumption, and stock returns. From this figure, it is clear that the Normal distribution is a poor description of the data because the series are, in fact, negatively skewed and leptokurtic. More formally, the hypothesis that the data follow a Normal distribution is tested using Jarque-Bera and Lilliefors tests and results are reported in table 3. The Jarque-Bera test is based on sample estimates of the skewness and excess kurtosis, both of which are should be zero if the data are normally distributed. Under the null hypothesis, the test statistic is distributed chi-square with two degrees of freedom. The Lilliefors test is a version of the Kolmogorov-Smirnov test designed to test the null hypothesis that the data is generated from a Normal distribution with unknown mean and standard deviation. The critical values for this test were tabulated by Lilliefors (1967). Panel A in table 3 reports results of tests applied to the data series used to estimate the model. Productivity innovations refers to the residuals of an autoregression for productivity with number of lags equal to 1, 2, and 6.[8] Note that all Jarque-Bera and Lilliefors test statistics are well above standard critical values and p-values are, correspondingly, small. Thus, the null hypothesis that the data is normally distributed can be rejected.

[7] In order to examine the robustness of results to the measure of consumption, I also estimated the model using personal consumption expenditures (PCE) divided by the PCE deflator and the civilian non-institutional population. Results are basically the same as those reported here and are available upon request.

[8] I performed tests on residuals of AR(1) through AR(12) processes but, in order to save space, the table only reports results based on AR(1), AR(2), and AR(6) processes. Results for the other autogressions are basically the same as those reported and are available upon request.

The robustness of this conclusion to the detrending procedure and the sample period is examined in panels B through D. Panel B reports results of tests applied to first-differenced, rather than detrended, data for the same sample period used to estimate the model. Panel C and D report results of tests respectively applied to detrended and first-differenced data for the longer sample from January 1959 to June 2014. As before, all test statistics are well above standard critical values and the null hypothesis that the data is normally distributed can be rejected.

The finding that productivity shocks are not well approximated by the Normal distribution is important because this is by far the most common distributional assumption in the business cycle literature. One exception is Curdia, del Negro, and Greenwald (2014), who model fat tails (but not asymmetry) using a t distribution and find that rare large shock realizations can generate "great recessions." This finding is also relevant for the macro-finance literature, which, rather than taking returns or consumption as given, attempts to push further the notion of fundamentals to the level of structural shocks like productivity. Asymmetry and fat tails in productivity shocks are, therefore, potentially important in asset pricing and potentially helpful in accounting for findings in earlier literature that documents a negative correlation between co-skewness with the marker portfolio and mean returns (Harvey and Siddique, 2000) and a negative effect of skewness risk on excess returns in a cross-section of stock returns and options (Kapadia, 2006, and Chang, Christoffersen, and Jacobs, 2012, respectively).

4.3 Implementation

The model is estimated using the indirect inference procedure proposed in section 3. That is, by finding the parameters that minimize the distance between the nonlinear impulse-responses generated by a nonlinear vector autoregression estimated using actual U.S. data and using data simulated from the model.

The time-series model for productivity is an AR(2) with innovations drawn from a Gumbel distribution with scale parameter ζ. As shown in figure 1, this distribution is negatively skewed and leptokurtic and, thus, describes the unconditional distribution of the innovations better than the Normal distribution. As benchmark, I also estimate a version of the model with Normal innovations. I use an AR(2) because this process is flexible and allows many possible shapes for the impulse responses, but it is also relatively parsimonious. Parsimony is important because, as discussed in section 2, productivity lags are states variables of the model and the computational cost of solving and simulating the model increases quickly with the size of the state vector. (For example, solving and simulating the model with 2 state variables takes 24 seconds while with 3 state variables it takes 30 seconds in a Dell XPS with an Core i7 processor.)

Based on arguments made in section 3.1, I use a NVAR(2,3) as auxiliary model—that is, a nonlinear vector autoregression with 2 lags and polynomial order equal to 3. The number of lags is that of the autoregressive process for productivity and the polynomial order is that of the perturbation. The binding function are nonlinear impulse-responses to productivity shocks in the 5th and 95th percentiles of the innovation distribution and propagated through the NVAR for a horizon of 50 periods. The weighting matrix (W) is the identity matrix and the variance-covariance matrix of the responses is computed using 500 simulations of the NVAR and its responses with shocks randomly drawn with replacement from the residuals.

4.4 Parameter Estimates

Table 2 reports estimates of the intertemporal elasticity of substitution (ψ), the coefficient of risk aversion (γ), the capital adjustment-cost parameter (ϕ), the autoregressive coefficients of the productivity shock (ρ_1 and ρ_2), and the standard deviation of the productivity innovations under the Gumbel and Normal distributions. In the former case, the standard deviation is $\pi\zeta/\sqrt{6}$, where ζ is the scale parameter and π is the mathematical constant. During the estimation procedure, the discount factor (β) was fixed to 0.9988, the depreciation rate was fixed to $\delta = 0.0075$, and the labor share ($1 - \alpha$) was fixed to 0.66. The motivation for these choices was explained in section 3.4.

Estimates of the intertemporal elasticity of substitution are 0.30 and 0.41 depending on whether innovations follow a Gumbel or a Normal distribution. These values are quantitatively close to estimates reported in earlier literature by Hall (1988), Epstein and Zin (1991) and Vissing-Jørgensen (2002): Hall reports estimates between 0.07 and 0.35, Epstein and Zin between 0.18 and 0.87 depending on the measure of consumption and instruments used, and Vissing-Jørgensen between 0.30 and 1 depending on the households' asset holdings. The coefficient of relative risk aversion are respectively 38.6 and 21.4 for the models with Gumbel and Normal productivity innovations. These estimates are smaller than the one of 79 reported by van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramirez (2012) and the values employed by calibration studies that use Epstein-Zin preferences. For example, Rudebusch and Swanson (2012) and Andreasen (2012) respectively employ 75 and 110. Finally, estimates of the productivity process are also very similar for both versions of the model and suggest that productivity is very persistent. We will see below that despite the similarity of parameter estimates, the models with Gumbel and Normal innovations generate very different dynamic predictions for consumption and stock returns.

4.5 Economic Implications

4.5.1 Higher-Order Moments

I first examine the model implications for two higher-order moments—skewness and kurtosis—and show that the version with asymmetric Gumbel innovations predicts moments that are generally closer to those of the data than those predicted by the version with symmetric Normal innovations. Column 1 of Table 5 reports the skewness and kurtosis of productivity innovations, consumption, and stock returns computed from the data. As we knew from figure 1, these series have negative skewness and positive excess kurtosis. Columns 2 and 3 respectively report the moments predicted by the model with Gumbel and Normal innovations, computed using simulated samples of 5000 observations. The version of the model with Normal innovations generates little skewness or excess kurtosis, and, in the case of stock returns, predicts skewness of sign opposite to the one in the data. This finding is non-trivial because the model is nonlinear and, hence, it could potentially generate skewness and excess kurtosis even if shocks are normally distributed. In contrast, the version of the model with Gumbel innovations generates substantial skewness although, in the case of consumption and stock returns, not as large as that of the data. This version also generates excess kurtosis for productivity innovations and stock returns, but not for consumption. Overall, I conclude that this version improves upon the one with Normal innovations in terms of higher-order moment predictions.

4.5.2 The Price of Risk

As pointed out above, policy functions make explicit the dependence of the model variables on state variables—capital and productivity—and on the moments of the productivity innovations. This is useful for quantifying the contribution of the different components of the solution. In particular, thinking of asset returns as the model variable, and with some slight abuse of the notation, the term $(1/2)[f_{\sigma\sigma}(s,0)]^j[\sigma][\sigma]$ consists of the price, $(1/2)[f_{\sigma\sigma}(s,0)]^j$, and quantity, $[\sigma][\sigma]$, of variance risk, while the term $(1/6)[f_{\sigma\sigma\sigma}(s,0)]^j[\sigma][\sigma][\sigma]$ consists of the price, $(1/6)[f_{\sigma\sigma\sigma}(s,0)]^j$, and quantity, $[\sigma][\sigma][\sigma]$, of skewness risk. Since the perturbating approximation was taken in logs, these prices have the interpretation of a semi-elasticity that measures the sensitivity of asset returns to changes in the quantity of risk. Table 6 reports estimates of the price of risk implied by the model at the indirect inference estimates when innovation follow a Gumbel and a Normal distribution. Notice that the price of variance risk is about the same for both distributions, but that the price of skewness risk is one order of magnitude larger than the price of variance risk.

Can this model explain the equity premium puzzle? No, and this is so despite the fact that this model features large risk aversion and a high capital adjustment costs. While the average real stock return in the data is about 0.55% per month (or 6.60% per year), the model with Gumbel innovations predicts a mean of only 0.12% per month (or 1.44% per year). The main reason for this result is that the overall quantity of risk in the model is relatively small. The estimated standard deviation of the productivity innovations is only 0.00095, which is much smaller than values used in calibrated macro-finance models. As pointed out by Campanale, Castro, and Clementi (2010), this poses a problem of the canonical growth model in terms of explaining asset returns. Related literature on disasters (e.g., see Andreasen, 2012, and Gourio, 2012)—which studies asset pricing in calibrated production economies using the shock formulation in Barro (2006), where innovations are drawn from the mixture of a normal and a Bernoulli distributions—reports higher mean returns but it treats productivity as a latent variable and so, it is unclear whether its time series properties are similar to those of Solow residuals estimated from the data.

4.5.3 Volatility Clustering

Most financial and macroeconomic series feature time-varying volatility and a large literature has developed in econometrics and finance to study this phenomenon. In this section, I show that consumption and stock returns, but not productivity, feature conditional heteroskedasticity and that the nonlinear model with Gumbel shocks can endogenously generate ARCH effects for stock returns.

Table 7 reports Lagrange Multiplier (LM) test statistics and p-values of the hypothesis of no conditional heteroskedasticity for the U.S. data. The test is carried out on the residuals of an autoregression with number of lags selected using the Akaike Information Criterion (AIC) and the statistic is calculated as the product of the number of observations and the uncentered R^2 of the OLS regression of squared residuals on a constant and two of its lags. Under the null hypothesis, the statistic is distributed chi-square with 2 degrees of freedom. The hypothesis can be safely rejected for consumption and stock returns at the monthly frequency but it cannot be rejected for productivity.

Table 7 also reports results for tests carried out on artificial data generated from the model with Gumbel and Normal shocks. As for the actual data, the hypothesis of no conditional heteroskedasticity cannot be rejected for productivity. Although the hypothesis cannot be rejected for consumption, it can be rejected for stock returns at the 10% significance level (p-value = 0.06) when shocks are drawn for the asymmetric Gumbel distribution. A key point is that this result arises despite the fact that shocks conditionally homoskedastic and, thus, illustrates the role of non-linearity as a mechanism that can endogenously generate predictable volatility clustering. The observation that nonlinear economic models can generate ARCH effects, even when shocks are i.i.d. and parameters are time-invariant, was first made by Granger and Machina (2006). Notice, however, that shock asymmetry also contributes to this result because for the version of the model where shocks are normally distributed, the hypothesis of no conditional heteroskedasticity cannot be rejected.

Finally, let us refer back to the generic formulation of the policy function in (11) and notice that it includes a time-varying term in the variance, $(1/2)[f_{x\sigma\sigma}(x,0)]^j_a[(x_t - x)]^a[\sigma][\sigma]$. This term makes the function resemble the ARCH-M model use by Engle et al. (1987) to study the term structure, where the conditional variance directly affects the mean. However, the key difference is that while in the ARCH-M model the conditional variance is time-varying and its coefficient is constant, in this model the conditional variance is constant (by assumption) and its coefficient is time-varying because it is a linear function of the state variables.

4.5.4 Impulse Responses

Figure 3 reports the impulses responses of consumption and stock returns to productivity innovations at the 5th and 95th percentile of the estimated Gumbel distribution (thick line). As a comparison, I also report the responses to innovations drawn from a Normal distribution with exactly the same standard deviation as the Gumbel distribution (thin line).

The asymmetry of the responses is clear in the figure. Responses to the (negative) shock in the 5th percentile are much larger effects in absolute value than the responses to the equally-likely (positive) shock in the 95th percentile. After the negative productivity shock, consumption declines steeply and reaches a minimum of 1.1% below trend after 9 months. Instead, after the positive shock, consumption increases moderately reaching only a maximum of 0.4% above trend, also after 9 months. In both cases, consumption returns to trend with a very modest overshooting. The dynamic effects of productivity shocks on stock returns are also asymmetric but they have a different pattern in that the effect is largest upon impact. The negative shock induces a drop of 0.8% in the period the shock takes place, while the positive shock induces an increase of 0.3%. (The units here are percentage deviations from the certainty-equivalent return). In contrast, if innovation are drawn from a Normal distribution, the model predicts similar quantitative effects for positive and negative shocks, basically as a result of the symmetry of the distribution.

An alternative way to appreciate the nonlinearity of the model is to plot the policy functions directly, as it is done in figure 4. The thick line is the relation between consumption (or stock returns) and productivity shocks implied by the nonlinear solution of the model at the parameter estimates for the Gumbel distribution reported in table 4. For comparison purposes, I plot the policy function under the linear solution at the same parameter estimates. The quantitative and qualitative importance of the nonlinearity is apparent from this figure. Overall, consumption and stock returns react by less to a positive productivity shock, but by much more to a negative productivity shock, than implied by a linear model. The reader may see this by comparing the distance between the thick and the thin lines for two shocks with the same magnitude but different sign. More generally, the quantitative response to a negative shock is much larger than the response to a positive shock of the same magnitude. The reader may see this by comparing the distance between the thick line and zero for two shocks with the same magnitude but different sign.

5. Conclusions

This paper proposes an impulse-response matching procedure explicitly designed to estimate nonlinear dynamic models and that employs as auxiliary model a new class of nonlinear vector autoregressions (NVAR). Under theoretically-based restrictions, the NVAR has a unique steady state and stable dynamics, and can be quickly estimated by least-squares methods. Monte-Carlo experiments show that indirect inference estimates that use the NVAR as auxiliary model are considerably more efficient than those that use a linear VAR.

The implementation of this procedure is illustrated by estimating a macro-finance model of asset pricing under skewness risk. Statistically, the motivating evidence for focusing on skewness risk are test results that show that productivity innovations are negatively skewed and that the hypothesis that the data are normally distributed is rejected by the data. Results show that the version of the model with asymmetric productivity innovations predicts higher-order moments that are qualitatively in line with the data, that the price of skewness risk is actually larger than the price of variance risk, that the nonlinear model can endogenously generate ARCH effects in asset returns, and that the responses of consumption and asset returns to productivity shocks are asymmetric, with negative shocks inducing larger responses than positive shocks.

In current and future work, I study further the properties of nonlinear vector autoregressions and their application to other problems in econometrics where nonlinearity is a key feature of the data or the model.

Table 1. Monte-Carlo Experiments

T=200

	ρ		η		γ		ϕ	
Model	True Median	Mean s.e.	True Median	Mean s.e.	True Median	Mean s.e.	True Median	Mean s.e.
Linear	0.9000	0.8961	0.1000	0.1004	2.0000	2.2193	100.00	105.71
	0.9001	0.0048	0.1003	0.0040	2.0290	0.1204	97.352	4.0005
Nonlinear	0.9000	0.8956	0.1000	0.1007	2.0000	2.1949	100.00	106.41
	0.8999	0.0025	0.1008	0.0020	2.0166	0.0647	99.050	2.2546
Linear	0.5000	0.5069	0.1000	0.1004	2.0000	2.0742	100.00	103.68
	0.5121	0.0106	0.1003	0.0041	2.0483	0.0827	103.93	4.0309
Nonlinear	0.5000	0.5005	0.1000	0.1025	2.0000	2.0519	100.00	117.04
	0.5112	0.0054	0.1025	0.0020	2.0678	0.0525	116.77	3.2597
Linear	0.9000	0.8961	0.1000	0.1004	20.000	25.000	100.00	109.73
	0.9001	0.0052	0.1003	0.0041	21.621	1.8590	101.14	5.6400
Nonlinear	0.9000	0.8960	0.1000	0.1006	20.000	23.987	100.00	110.95
	0.9000	0.0025	0.1006	0.0020	20.989	0.8728	100.35	2.9300
Linear	0.5000	0.5069	0.1000	0.1004	20.000	21.436	100.00	106.97
	0.5121	0.0105	0.1003	0.0042	21.180	0.8935	106.94	2.8024
Nonlinear	0.5000	0.4924	0.1000	0.1049	20.000	20.433	100.00	118.33
	0.5017	0.0060	0.1042	0.0023	20.152	0.5358	116.88	2.7753
Linear	0.9000	0.8961	0.0100	0.0100	2.0000	2.2139	100.00	104.65
	0.9001	0.0048	0.0100	0.0004	2.0344	0.1235	97.214	3.9889
Nonlinear	0.9000	0.8958	0.0100	0.0101	2.0000	2.1881	100.00	105.22
	0.8999	0.0025	0.0101	0.0002	2.0118	0.0656	98.054	2.1856
Linear	0.5000	0.5069	0.0100	0.0100	2.0000	2.0697	100.00	102.43
	0.5121	0.0104	0.0100	0.0004	2.0473	0.0847	102.22	4.0157
Nonlinear	0.5000	0.5010	0.0100	0.0102	2.0000	2.0465	100.00	114.23
	0.5118	0.0053	0.0102	0.0002	2.0625	0.0514	113.11	3.0604
Linear	0.9000	0.8961	0.0100	0.0104	20.000	24.607	100.00	106.42
	0.9001	0.0049	0.0100	0.0004	21.265	1.6994	99.079	4.8634
Nonlinear	0.9000	0.8960	0.0100	0.0101	20.000	23.572	100.00	106.76
	0.9000	0.0024	0.0101	0.0002	20.905	0.8174	98.817	2.4575
Linear	0.5000	0.5069	0.0100	0.0100	20.000	21.185	100.00	102.98
	0.5121	0.0109	0.0100	0.0004	20.947	0.8776	102.81	2.3617
Nonlinear	0.5000	0.4995	0.0100	0.0103	20.000	21.035	100.00	107.04
	0.5105	0.0057	0.0103	0.0002	20.750	0.5132	110.10	1.5334

Note: True is the parameter value used to generate the data, Mean and Median are, respectively, the mean and the median of the estimated parameter values; and s.e. is the mean asymptotic standard error. Linear refers to the model estimated using a VAR(1). Nonlinear refers to the model estimated using a NVAR(1,2). The model was solved using a second-order perturbation.

Table 2. Monte-Carlo Experiments

T=500

	ρ		η		γ		ϕ	
Model	True Median	Mean s.e.	True Median	Mean s.e.	True Median	Mean s.e.	True Median	Mean s.e.
Linear	0.9000	0.8954	0.1000	0.1000	2.0000	2.1006	100.00	94.740
	0.8969	0.0025	0.1001	0.0018	2.0240	0.0590	92.532	1.6483
Nonlinear	0.9000	0.8952	0.1000	0.1002	2.0000	2.0686	100.00	94.982
	0.8967	0.0012	0.1003	0.0009	1.9818	0.0291	92.390	0.8647
Linear	0.5000	0.4955	0.1000	0.1000	2.0000	2.0232	100.00	98.783
	0.4932	0.0051	0.1002	0.0016	2.0020	0.0386	97.851	1.7445
Nonlinear	0.5000	0.4962	0.1000	0.0999	2.0000	2.0169	100.00	97.570
	0.4954	0.0026	0.1001	0.0008	1.9846	0.0239	98.303	1.0441
Linear	0.9000	0.8954	0.1000	0.1000	20.000	33.085	100.00	94.963
	0.8969	0.0025	0.1001	0.0017	29.952	1.5003	91.782	1.9808
Nonlinear	0.9000	0.8954	0.1000	0.1001	20.000	30.150	100.00	95.178
	0.8968	0.0012	0.1002	0.0008	26.656	0.6397	92.989	1.0104
Linear	0.5000	0.4955	0.1000	0.1000	20.000	22.901	100.00	99.655
	0.4932	0.0051	0.1002	0.0018	22.511	0.4942	99.588	1.1528
Nonlinear	0.5000	0.4967	0.1000	0.0998	20.000	22.098	100.00	98.733
	0.4955	0.0027	0.0999	0.0009	21.798	0.2854	99.570	0.8360
Linear	0.9000	0.8954	0.0100	0.0100	2.0000	2.1018	100.00	94.821
	0.8969	0.0025	0.0100	0.0002	2.0279	0.0594	93.054	1.6755
Nonlinear	0.9000	0.8952	0.0100	0.0100	2.0000	2.0689	100.00	95.116
	0.8966	0.0012	0.0100	0.0001	1.9866	0.0287	93.378	0.8511
Linear	0.5000	0.4955	0.0100	0.0100	2.0000	2.0222	100.00	98.430
	0.4932	0.0051	0.0100	0.0002	2.0042	0.0380	97.432	1.7096
Nonlinear	0.5000	0.5010	0.0100	0.0102	2.0000	2.0465	100.00	114.23
	0.5118	0.0053	0.0102	0.0002	2.0625	0.0514	113.11	3.0604
Linear	0.9000	0.8954	0.0100	0.0100	20.000	33.209	100.00	95.287
	0.8969	0.0024	0.0100	0.0002	30.131	1.4418	93.201	1.8862
Nonlinear	0.9000	0.8953	0.0100	0.0100	20.000	30.257	100.00	95.551
	0.8968	0.0012	0.0100	0.0001	26.427	0.6461	93.650	0.9793
Linear	0.5000	0.4955	0.0100	0.0100	20.000	22.842	100.00	98.596
	0.4932	0.0384	0.0100	0.0003	22.509	1.9792	97.786	7.6314
Nonlinear	0.5000	0.4967	0.0100	0.0100	20.000	21.983	100.00	97.702
	0.4953	0.0026	0.0100	0.0001	21.701	0.2677	98.686	0.6173

Note: See notes to table 1.

Table 3. Normality Tests

	Jarque-Bera		Lilliefors	
	Statistic	p-value	Statistic	p-value
	A. Data Used to Estimate the Model			
Productivity	15.00	0.004	0.062	< 0.001
Productivity innovations				
AR(1)	46.35	< 0.001	0.062	< 0.001
AR(2)	32.65	< 0.001	0.067	< 0.001
AR(6)	17.23	0.002	0.096	< 0.001
Consumption	45.07	< 0.001	0.099	< 0.001
Stock returns	256.8	< 0.001	0.054	< 0.001
	B. Growth Rates			
Productivity growth	61.12	< 0.001	0.064	< 0.001
Productivity innovations				
AR(1)	111.9	< 0.001	0.090	< 0.001
AR(2)	102.3	< 0.001	0.104	< 0.001
AR(6)	88.69	< 0.001	0.120	< 0.001
Consumption growth	324.3	< 0.001	0.070	< 0.001
	C. Longer Sample and Detrended Data			
Productivity	12.45	0.006	0.060	< 0.001
Productivity innovations				
AR(1)	59.90	< 0.001	0.059	< 0.001
AR(2)	28.87	< 0.001	0.049	< 0.001
AR(6)	9.871	0.012	0.076	< 0.001
Consumption	16.94	0.002	0.057	< 0.001
	D. Longer Sample and Growth Rates			
Productivity growth	54.27	< 0.001	0.065	< 0.001
Productivity innovations				
AR(1)	305.7	< 0.001	0.064	< 0.001
AR(2)	290.3	< 0.001	0.093	< 0.001
AR(6)	174.9	< 0.001	0.118	< 0.001
Consumption growth	220.9	< 0.001	0.063	< 0.001

Note: Under the null hypothesis that the data follow a Normal distribution, the statistic of the Jarque-Bera test is distributed chi-square with two degrees of freedom. The critical values for the Lilliefors test are tabulated in Lilliefors (1967).

Table 4. Parameter Estimates

Parameter	Distribution			
	Gumbel		Normal	
	Estimate	s.e.	Estimate	s.e.
IES	0.3021*	0.0036	0.4129*	0.0047
Risk aversion	38.611*	15.795	21.400	15.737
Capital adjustment ($\times 10^3$)	81.545	167.87	163.49	1521.1
AR coefficient	1.7898*	0.0009	1.7992*	0.0010
AR coefficient	−0.8056*	0.0008	−0.8128*	0.0009
Scale parameter ($\times 10^{-3}$)	0.7421*	0.0038	–	–
Standard deviation ($\times 10^{-3}$)	0.9517*	0.0048	0.8202*	0.0066

Note: The superscript $*$ denotes statistical significance at the 5% level.

Table 5: Higher-Order Moments

	U.S.	Distribution	
	Data	Gumbel	Normal
Productivity innovations			
Skewness	−0.206	−1.031	−0.046
Kurtosis	5.346	4.689	2.991
Consumption			
Skewness	−0.719	−0.197	−0.116
Kurtosis	3.057	2.954	3.083
Stock returns			
Skewness	−0.827	−0.355	0.076
Kurtosis	6.012	3.533	2.954

Note: The table reports unconditional moments of actual U.S. series and of artificial data simulated from estimated models. The sample size of the artificial data is 5000 observations.

Table 6. Price of Risk

	Distribution	
Source	Gumbel	Normal
Variance	4.187	4.123
Skewness	53.31	–

Notes: The table reports the price of different risk sources under each of the statistical distributions considered.

Table 7. Test of No Conditional Heteroskedasticity

	U.S. Data		Distribution: Gumbel		Distribution: Normal	
	Statistic	p-value	Statistic	p-value	Statistic	p-value
Productivity	1.043	0.594	2.499	0.287	0.712	0.701
Consumption	8.095	0.018	2.542	0.453	0.732	0.693
Stock returns	6.900	0.033	5.492	0.064	0.543	0.762

Note: Under the null hypothesis of no conditional heteroskedasticity, the statistic follows a chi-square distribution with two degrees of freedom. Tests on data generated from the model were carried out using artificial series with 5000 observations.

References

[1] Andreasen, M. M., 2012. On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-linear DSGE Models. Review of Economic Dynamics 15, pp. 295-316.

[2] Altig, D., Christiano, L., Eichenbaum, M., Linde, J., 2011. Firm Specific Capital, Nominal Rigidities and the Business Cycle. Review of Economic Dynamics 14, pp. 225-247.

[3] Aruoba, B., Bocola, L., Schorfheide, F., 2014. Assessing DSGE Model Nonlinearities. Manuscript.

[4] Barnichon, R, Matthes, C., 2014. Measuring the Non-Linear Effects of Monetary Policy. Manuscript.

[5] Barro, R. J., 2006. Rare Disasters and Asset Markets in the Twentieth Century. Quarterly Journal of Economics 121, pp. 823-866.

[6] Boivin, J., Giannoni, M., 2006. Has Monetary Policy Become More Effective? Review of Economics and Statistics 88, pp. 445-462.

[7] Caldera, D., Fernández-Villaverde, J., Rubio-Ramírez, J. F., 2012. Computing DSGE Models with Recursive Preferences. Review of Economic Dynamics 15, pp. 188–206.

[8] Campanale, C., Castro, R., and Clementi, G. L., 2010. Asset Pricing in a Production Economy with Chew-Dekel Preferences. Review of Economic Dynamics 13, pp. 379-402l.

[9] Canova, F., Sala, L., 2009. Back to Square One: Identification Issues in DSGE Models. Journal of Monetary Economics 56, pp. 431-449.

[10] Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy 113, pp. 1-45.

[11] Cúrdia, V., del Negro, M., Greenwald, D., 2014. Rare Shocks, Great Recessions. Journal of Applied Econometrics, forthcoming.

[12] DiCecio, R., Nelson, E., 2007. An estimated DSGE Model for the United Kingdom. Federal Reserve Bank of St. Louis Review 89, pp. 215-231.

[13] Dridi, R. Guay, A., Renault, E., 2007. Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models. Journal of Econometrics 136, pp. 397-430.

[14] Dupor, B., Han, J., Tsai, Y.-C., 2007. What Do Technology Shocks Tell Us About the New Keynesian Paradigm? Journal of Monetary Economics 56, pp. 560-569.

[15] Engle, R. F., Lilien, D. M., Robins., R. P., 1987. Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. Econometrica 55, pp. 391-407.

[16] Epstein, L., Zin, S., 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica 57, pp. 937-969.

[17] Epstein, L., Zin, S., 1991. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns. Journal of Political Economy 99, pp. 263-286.

[18] Fernald, J. G., 2014. A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Federal Resever Bank of San Francisco Working Paper 2012-19.

[19] Gallant, A. R., Rossi, P. E., Tauchen, G., 1993. Nonlinear Dynamic Structures. Econometrica 61, pp. 871-908.

[20] Granger, C., Machina, M., 2006. Structural Attribution of Observed Volatility Clustering. Journal of Econometrics 135, pp. 15–29.

[21] Gourieroux, C., Monfort, A., and Renault, E., 1993. Indirect Inference. Journal of Applied Econometrics 8, S85-S118.

[22] Gumbel, E. J. 1935. Les valeurs extrêmes des distributions statistiques. Annales de l'Institute Henri Poincaré 5, pp. 115-158.

[23] Gurio, F., 2012, Disaster Risk and Business Cycles. American Economic Review 102, pp. 2734–2766.

[24] Hall, R. E., 1988. Intertemporal Substitution in Consumption. Journal of Political Economy 96, pp. 339-357.

[25] Hall, A., Inoue, A., Nason, J., Rossi, B., 2012. Information Criteria for Impulse Response Function Matching Estimation of DSGE Models. Journal of Econometrics 170, pp. 499-518.

[26] Hansen, L. P., Singleton, K., 1982. Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models. Econometrica 50, pp. 1269-1286.

[27] Harvey, C. R., Siddique, A., 2000. Conditional Skewness and Asset Pricing Tests. Journal of Finance 55, pp. 1263-1295.

[28] Iacoviello, M., 2005. House Prices, Borrowing Constraints and Monetary Policy in the Business Cycle. American Economic Review 95, pp. 739-764.

[29] Jermann, U. J., 1998. Asset Pricing in Production Economies. Journal of Monetary Economics 41, pp. 257-275.

[30] Jin, H., Judd, K. L., 2002. Perturbation methods for general dynamic stochastic models. Hoover Institution, Mimeo.

[31] Jorda, O., Kozicki, S., 2007. Estimation and Inference by the Method of Projection Minimum Distance, Bank of Canada Working Paper 2007-56.

[32] Kapadia, N., 2006. The Next Microsoft? Skewness, Idyosincratic Volatility, and Expected Returns. Rice University, Mimeo.

[33] Kim. J., Kim, S., Schaumburg E., Sims, C., 2008. Calculating and Using Second-Order Accurate Solutions of Discrete-Time Dynamic Equilibrium Models. Journal of Economic Dynamics and Control 32, pp. 3397-3414.

[34] Koop, G., Pesaran, M. H., Potter, S., 1996. Impulse Response Analysis in Nonlinear Multivariate Models. Journal of Econometrics 74, pp. 119-147.

[35] Lilliefors, H. W., 1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. Journal of the American Statistical Association 62, pp. 399–402.

[36] Mandelbrot, B., Hudson, R., 2004. The (mis)Behavior of Markets. New York: Basic Books.

[37] Mittnik, S., 1990. Modeling Nonlinear Processes With Generalized Autoregressions. Applied Mathematics Letters 3, pp. 71-74.

[38] Rotemberg, J. Woodford, M., 1998. An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. In: Bernanke, B., Rotemberg, J. (Eds.), NBER Macroeconomics Annual. MIT Press: Cambridge, MA.

[39] Rudebusch, G., Swanson, E., 2012. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. American Economic Journal: Macroeconomics 4, pp. 105-143.

[40] Ruge-Murcia, F., 2012. Estimating Nonlinear DSGE Models by the Simulated Method of Moments: With an Application to Business Cycles, Journal of Economic Dynamics and Control 36, pp. 914-938.

[41] Ruge-Murcia, F., 2014. Skewness Risk and Bond Prices, McGill University, Mimeo.

[42] Schmitt-Grohé, S., Uribe, M., 2004. Solving dynamic general equilibrium models using a second-order approximation to the policy function. Journal of Economic Dynamics and Control 28, pp. 755-775.

[43] Smith, A., 1993. Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions. Journal of Applied Econometrics 8, S63-S84.

[44] Smith, A., 2008. Indirect Inference. In: The New Palgrave Dictionary of Economics. Palgrave Macmillan: Basingstoke.

[45] Uribe, M., Yue, V., 2006. Country Spreads and Emerging Markets: Who Drives Whom? Journal of International Economics 69, pp. 6-36.

[46] van Binsbergen, J., Fernández-Villaverde, J., Koijen, R, Rubio-Ramirez, J., 2012. The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences. Journal of Monetary Economics 59, pp. 634-648.

[47] Varian, H., 1974. A Bayesian Approach to Real Estate Assessment. In: Feinberg S. E., Zellner, A., (Eds.), Studies in Bayesian Economics in Honour of L. J. Savage. North-Holland, Amsterdam.

[48] Vissing-Jørgensen, A., 2002. Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. Journal of Political Economy 110, pp. 825-853.

Figure 1: Probability Density Functions

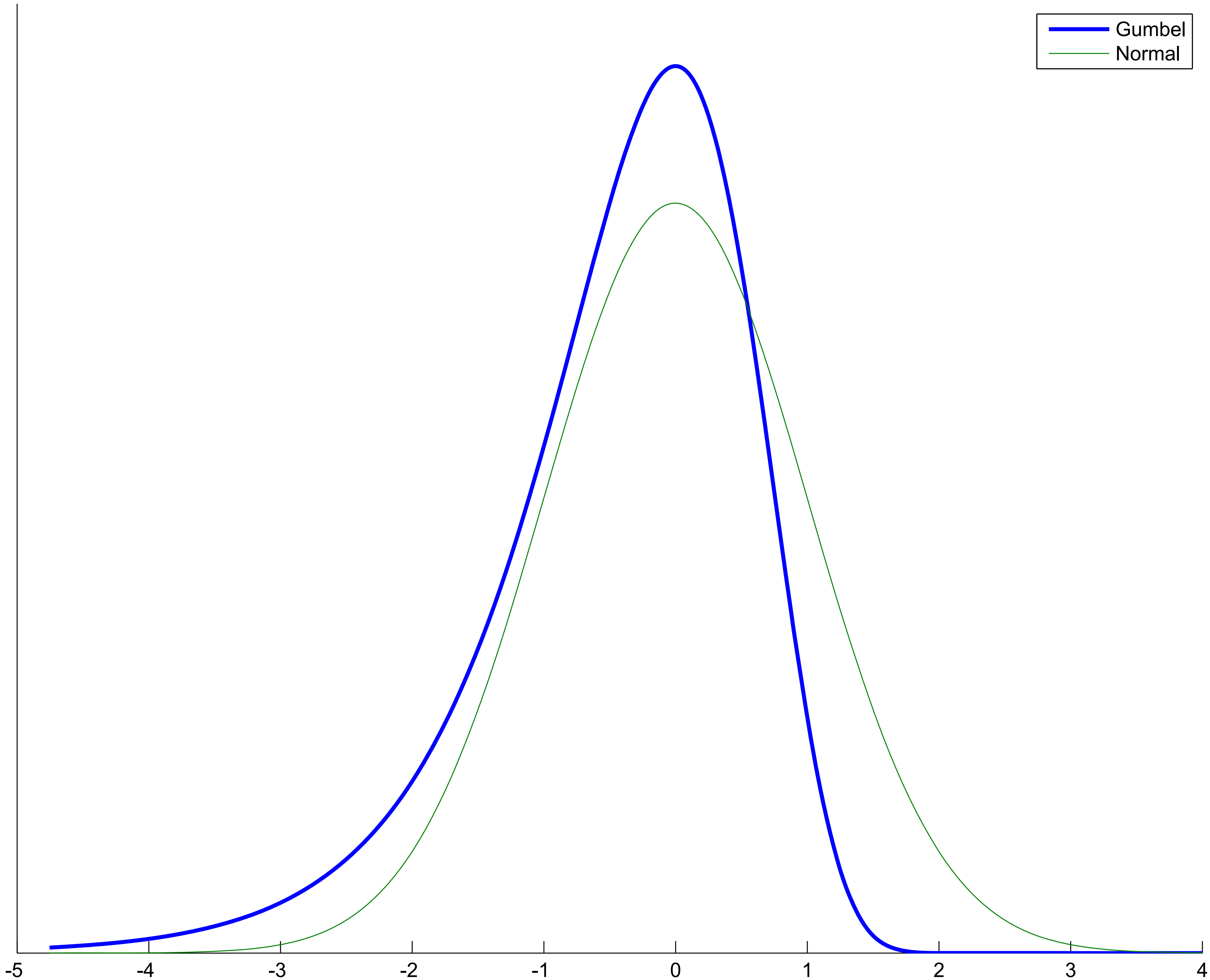


Figure 2: Asymmetry in U.S. Data

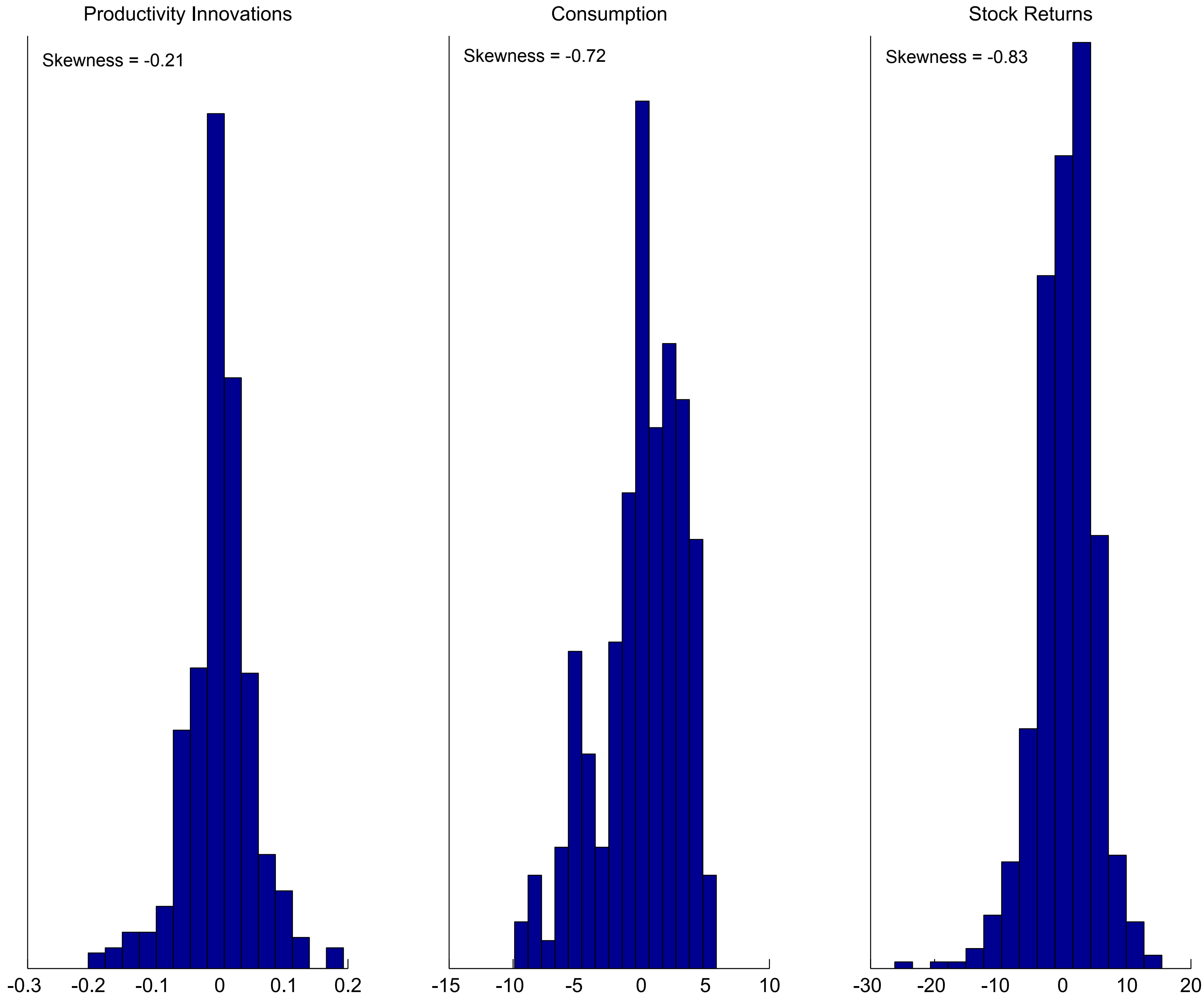


Figure 3: Impulse Responses to a Productivity Shock

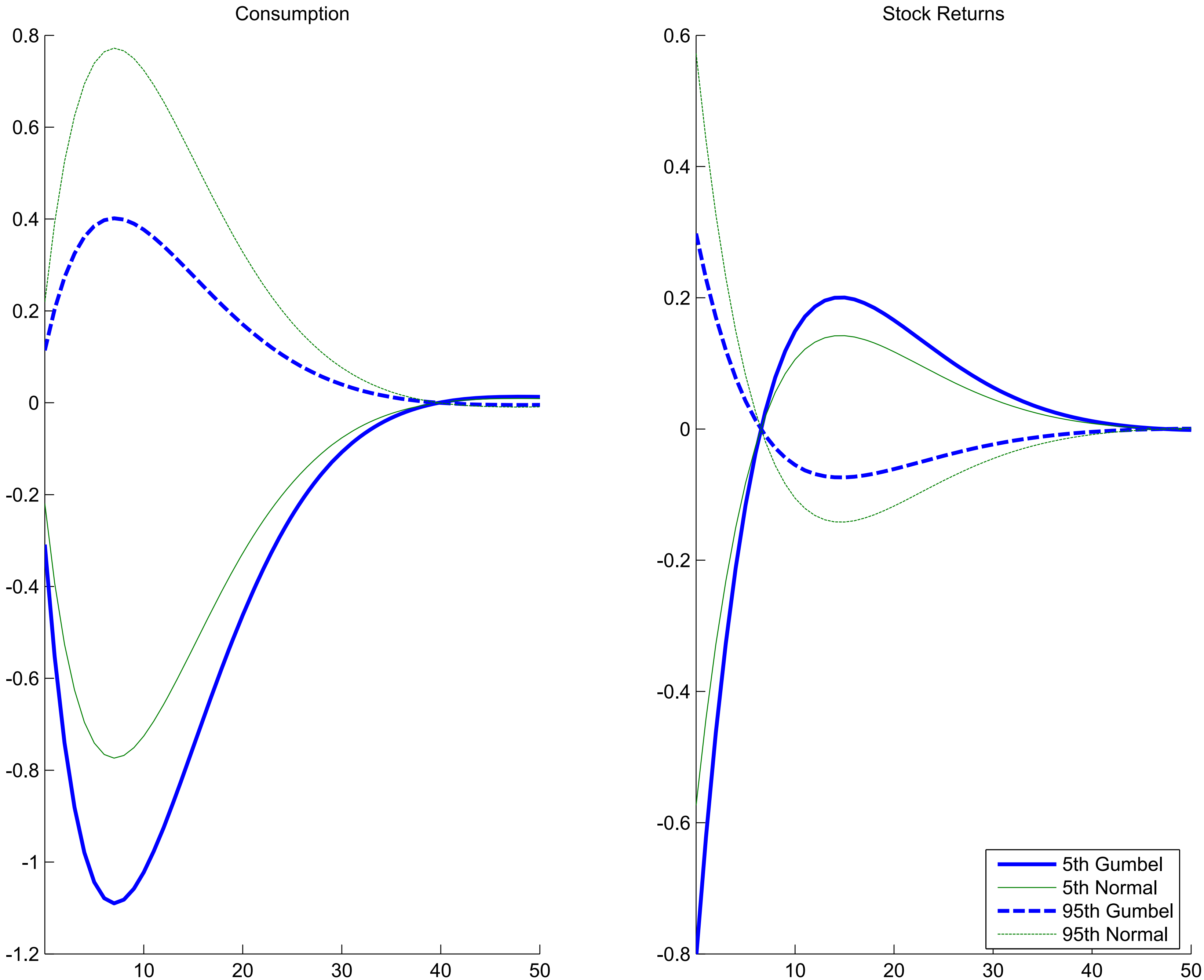
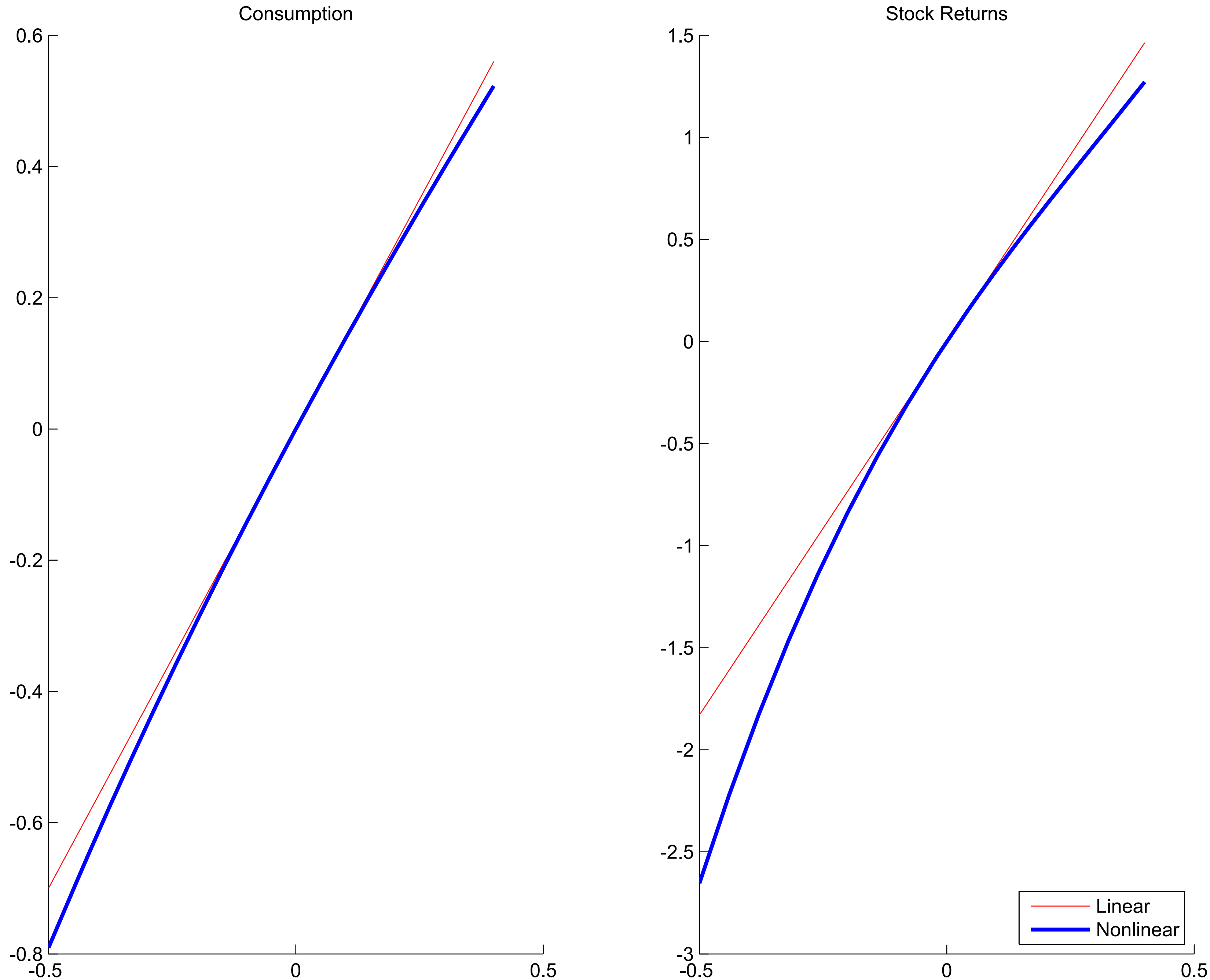


Figure 4: Relation with Productivity Shocks



Récents cahiers de recherche du CIREQ
Recent Working Papers of CIREQ

Si vous désirez obtenir des exemplaires des cahiers, vous pouvez les télécharger à partir de notre site Web http://www.cireqmontreal.com/cahiers-de-recherche

If you wish to obtain copies of the working papers, you can download them directly from our website, http://www.cireqmontreal.com/cahiers-de-recherche

01-2014	Alvarez-Cuadrado,F., N.V. Long et M. Poschke, "Capital-Labor Substitution, Structural Change and the Labor Income Share", janvier 2014, 55 pages
02-2014	Long, N.V., "The Green Paradox under Imperfect Substitutability between Clean and Dirty Fuels", février 2014, 23 pages
03-2014	Doko Tchatoka, F., J.-M. Dufour, "Identification-Robust Inference for Endogeneity Parameters in Linear Structural Models", janvier 2014, 29 pages
04-2014	Amarante, M., "What Is Ambiguity?", mars 2014, 51 pages
05-2014	Ehlers, L., "Strategy-Proofness and Essentially Single-Valued Cores Revisited", avril 2014, 22 pages
06-2014	Laussel, D., N.V. Long, J. Resende, "Network Effects, Aftermarkets and the Coase Conjecture : A Dynamic Markovian Approach", mai 2014, 65 pages
07-2014	Bossert, W., K. Suzumura, "Expected Utility without Full Transitivity", mai 2014, 12 pages
08-2014	Gaudet, G., S.W. Salant, "The Hotelling Model with Multiple Demands", octobre 2013, 22 pages
09-2014	Andersson, T., L. Ehlers, L.-G. Svensson, "Transferring Ownership of Public Housing to Existing Tenants : A Mechanism Design Approach", août 2014, 30 pages
10-2014	Negri, M., Y. Sprumont, "Size Invariant Measures of Association : Characterization and Difficulties", août 2014, 26 pages
05R-2011	Castro, R., N. Koumtingué, "On the Individual Optimality of Economic Integration", août 2014, 54 pages
11-2014	Castro, R., D. Coen-Pirani, "Explaining the Evolution of Educational Attainment in the U.S.", juillet 2014, 64 pages
15R-2010	Castro, R., G.L. Clementi, Y. Lee, "Cross-Sectoral Variation in the Volatility of Plant-Level Idiosyncratic Shocks", juin 2013, 33 pages
12-2014	Benchekroun, H. S. Benchekroun, "Harvests' Lifespan and North-South Market Share Rivalry", octobre 2014, 20 pages
13-2014	Daubanes, J., P. Lasserre, "Dispatching after Producing : The Supply of Non-Renewable Resources", octobre 2014, 63 pages
14-2014	Ehlers, L, B. Klaus, "Object Allocation via Deferred-Acceptance : Strategy-Proofness and Comparative Statics", décembre 2014, 30 pages