Dynamics of Externalities: A Second-Order Perspective

Yi Wen
and
Huabin Wu

Working Paper 2008-044B

November 2008
Revised March 2011

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.
Abstract

First-order approximation methods are a standard technique for analyzing the local dynamics of dynamic stochastic general equilibrium (DSGE) models. Although for a wide class of DSGE models linear methods yield quite accurate solutions, some important economic issues such as portfolio choice and welfare cannot be adequately addressed by first-order methods. This paper provides yet another case where first-order methods may be inadequate for capturing a DSGE model’s business-cycle properties. In particular, we show that increasing returns to scale (due to production externalities) may induce asymmetric business cycles and nonlinear income effects that are not fully appreciated by linear approximation methods. For example, hump-shaped output dynamics can emerge even when externalities are below the threshold level required for indeterminacy, and output expansion tends to be smoother and longer while contraction tends to be deeper but shorter-lived, as observed in the U.S. economy.

JEL classification: C63; E0; E32.

Keywords: Second-order approximation; Asymmetric business cycle; Hump-shaped output dynamics; Increasing returns to scale.

*We thank Adrian Peralta-Alva, Massimo Guidolin, and Bill Gavin for comments.
1 Introduction

The standard approach to studying the business-cycle implications of DSGE models is to focus on the model’s local dynamics near a steady state through linear (first-order) approximations (such as the loglinearization method of King, Plosser, and Rebelo, 1988). It is well-known that for standard real business cycle (RBC) models with constant returns, first-order approximation methods give quite accurate solutions and higher-order methods give almost identical predictions.

The central twist of this paper is the addition of increasing returns to scale (IRS) due to production externalities. This simple deviation from standard RBC models is shown to generate nontrivial non-linearities that are not well captured by first-order methods. Importantly, these non-linearities are increasing in the degree of external economies over parameter ranges that predict a unique bounded rational expectations equilibrium. The model does not rely on local indeterminacy to generate new and interesting dynamics, though one contribution of the paper is to document model properties when the model gives rise to local indeterminacy.¹

Specific results of interest are that technology shocks generate asymmetric effects on business cycles. These effects are generated by second-order components of the model, which are ignored by linear approximation methods. Conditional on a positive technology shock, hump-shaped impulse response functions are predicted for employment and output, consistent with much empirical work. Smooth, prolonged dynamics are observed. In contrast, conditional on a negative technology shock, the model predicts sharp, less persistent dynamics. Combined these insights provide an explanation of the strong asymmetry of the business cycle in the U.S. economy.

IRS have been shown in the existing literature as an important source of dynamics not only for endogenous growth (e.g., Romer, 1986) but also for the business cycle. Baxter and King (1991) show that incorporating production externalities into a standard RBC model generates a better overall fit of the model to the U.S. data, especially under aggregate demand shocks. Benhabib and Farmer (1994), Farmer and Guo (1994), Wen (1998a), and Benhabib and Wen (2004) show that IRS can generate endogenous business cycles if externalities are large enough to make the model’s steady state locally indeterminate.²

¹Local indeterminacy means that there are multiple rational-expectations equilibrium paths that converge to the same steady state. Note that local indeterminacy is not the same thing as multiple steady-state equilibria. Indeterminacy can arise in a model with a unique steady state.

²Also see Cooper and Johri (1997) and Wen (1998b) for business-cycle implications of externalities in RBC models without indeterminacy.
However, this segment of the literature has based on first-order approximations to investigate the model’s dynamic behaviors. Although it has been shown that, for standard RBC models (such as the models of Kydland and Prescott, 1982; and King, Plosser, and Rebelo, 1988), first-order approximation usually yields quite accurate results, it is less clear whether such methods remain accurate in describing a model’s local dynamics when it involves market failures and non-convexities, such as externalities and IRS.

IRS can greatly amplify the impact of shocks (Baxter and King, 1991) and dramatically change a model’s topology around its steady state so that complicated dynamics, such as bifurcations, discontinuous jumps, and complex eigenvalues, may emerge (Benhabib and Farmer, 1994; and Coury and Wen, 2009). Therefore, it is interesting to investigate whether first-order approximation methods continue to yield accurate predictions of a model’s dynamics when there are production externalities.

In addition, one of the most important aspects of stochastic dynamic models—risk—cannot be well captured by linear solution methods. As an example, optimal portfolio decisions can not be analyzed under linear approximation methods. For reasons like this, second-order solution methods have been proposed and developed by the recent literature. For a review of this literature, see, for example, Judd (1998), Jin and Judd (2002), Collard and Juillard (2001), Schmitt-Grohe and Uribe (2004), Anderson et al. (2006), Lombardo and Sutherland (2007), and Kim et. al. (2008), among others.

In this paper, we apply the second-order approximation method developed by the existing literature to analyze the local dynamics of an RBC model with externalities. We show that taking second-order terms into consideration not only improves the accuracy of approximations but also changes the predicted local dynamics of the model dramatically when IRS exist. In particular, the magnitude of the impulse responses to positive technology shocks are significantly smaller and smoother under second-order approximation method than those found under the first-order method, and hump shaped impulse responses can emerge even with degrees of externalities that are too small to trigger indeterminacy. This is in sharp contrast to the results in Benhabib and Wen (2004) where, under the first-order approximation method, hump-shaped dynamics emerge only when the degree of externalities is large enough so that the model becomes locally indeterminate.

The new findings regarding the dynamic effects of externalities are driven by the fact that externalities introduce a large nonlinear income effect—captured by a large negative coefficient in front of the square term of technology shocks and a large positive coefficient in front of the cross term of capital and technology shock. This implies that technology
shocks have the following asymmetric second-order effects on aggregate output dynamics: A positive technology shock generates smooth and hump-shaped output responses because the variance of technology shocks neutralize the shock on impact but the covariance of capital and technology shocks enhances (and thus propagates) the effects of the positive shock in the subsequent periods. On the other hand, a negative technology shock generates much bigger and sharper falls in output relative to the first-order method because the variance of technology shocks greatly amplify the negative shock while the covariance term reduces its negative impact in the subsequent periods, making the impulse responses of output monotonically increasing towards the steady state from below. Such an asymmetric nonlinear effect is similar to what we observe in the U.S. economy: on the one hand, economic expansion tends to be smooth and gradual while economic contraction tends to be sharp and short lived (see, e.g., Neftic, 1984; Sichel, 1993; Kim and Piger, 2002; McKay and Reis, 2008; and Morley and Piger 2009); on the other hand, regarding the effects of oil shocks, unexpected increases in oil prices (equivalent to negative TFP shocks) tend to have large adverse effect on aggregate output while decreases in oil prices tend to have only small or negligible positive effects on output (e.g., see Hamilton, 2003; and Mork, 1989).

Our analysis also helps to explain a puzzle in the indeterminacy literature. That is (e.g., in the Benhabib-Farmer model), under the first-order approximation method, positive technology shocks are extremely expansionary before indeterminacy arises but suddenly become excessively contractionary when the model becomes locally indeterminate, generating sharp falls in output, investment, and hours. This is puzzling because it indicates an unusually large income effect on hours worked and savings not observed in standard RBC models with constant returns. Using second-order methods, we find that positive technology shocks have a second-order negative impact on hours worked and this nonlinear income effect increases with the degree of increasing returns to scale. Hence, as externalities increase, the second-order income effect gradually dominates the first-order substitution effect, and the initial response of hours gradually turns to negative around the point of indeterminacy. Once the model becomes indeterminate, the coefficients of second-order terms in labor become zero so that the strong income effect are captured only by the first-order terms.\(^3\)

An important caveat is that this paper’s main point is not to find frictions to improve the empirical fit of existing RBC models for the business cycle. Hence, moment matching is not our goal. Rather, we try to provide a new case where linear solution methods may give inaccurate descriptions about the model’s dynamics. Moreover, given that the existing

\(^3\)See discussions in section 3.3.
literature has shown that IRS may be important for understanding the business cycle based on first-order approximations, we are also interested in understanding whether IRS can introduce non-linear dynamics that have not been captured by first-order methods.

The point that linear approximation methods may result in quantitatively significant biases even for standard RBC models without externalities has been made in the literature. For example, Fernandez-Villaverde et al. (2006) study the econometrics of computed dynamic models and the consequences for inference of the use of approximated likelihoods. They find that second order approximation errors in the policy function, which are completely ignored by first-order approximation methods, have first order effects on the likelihood function.\(^4\)

The problem of numerical simulations of dynamic economies with heterogeneous agents and economic distortions are also studied by the existing literature. For example, Feng et al. (2009) and Peralta-Alva and Santos (2010), among others, discuss problems related to the existence and computation of Markovian equilibria in economies with heterogeneous agents and market distortions, as well as convergence and accuracy properties of numerical solutions.\(^5\)

This rest of the paper is organized as follows: Section 2 presents the model with externalities and introduces the second-order perturbation method. A simple example is provided to illustrate the difference between first-order and second-order methods. Section 3 examines the model’s dynamics with and without indeterminacy. Section 4 compares the accuracy of linear and second-order approximations in the presence of IRS, and Section 5 concludes the paper.

2 The Model

The model is similar to that of Baxter and King (1991) and Benhabib and Farmer (1994).\(^6\)

There exist a continuum of identical agents in the unit interval \([0,1]\). A typical or representative agent chooses consumption \((c_t)\), hours worked \((n_t)\), and capital stock \((k_{t+1})\) to

\(^4\)However, Ackerberg et al. (2009) show by counterexample that this conclusion of Fernández-Villaverde et al. (2006) is false and argue that second order approximation errors in the policy function have at most second order effects on parameter estimates.

\(^5\)More references to this literature can also be found in these two articles.

\(^6\)We have also studied the model of Wen (1998a) where the degree of externalities required for indeterminacy is much smaller because of variable capacity utilization, and our conclusions remain robust. To simplify the analysis, we choose the simpler model of Benhabib and Farmer (1998) with fixed capacity utilization even though this model requires an implausibly large degree of IRS to generate a significant difference in local dynamics from those of standard RBC models.
solve

$$\max_{\{c_t, n_t, k_{t+1}\}} E_t \sum_{t=0}^{\infty} \beta^t \left( \frac{c_t^{1-\tau} - a_t n_t^{1+\gamma}}{1 + \gamma} \right),$$

subject to the resource constraint, $c_t + k_{t+1} - (1 - \delta) \bar{k}_t \leq A_t (\bar{k}_t^{\alpha} \bar{n}_t^{1-\alpha})^{\eta_{t}} n_t^{1-\alpha}$, where $\{\bar{k}_t, \bar{n}_t\}$ denote the average economy-wide capital stock and hours which are taken as given by individuals and $A_t$ denotes aggregate technology shocks. The model exhibits increasing return to scale at the social level if the externality parameter $\eta > 0$.\(^7\)

The equilibrium of the model is determined by the following necessary conditions:

$$an_t^\gamma = c_{t+1}^{-\tau} (1 - \alpha) A_t k_{t+1}^{\alpha(1+\eta)} n_t^{(1-\alpha)(1+\eta)-1}$$ (1)

$$c_t^{-\tau} = \beta E_t \left\{ c_{t+1}^{-\tau} (\alpha A_{t+1} k_{t+1}^{\alpha(1+\eta)-1} n_{t+1}^{(1-\alpha)(1+\eta)} + 1 - \delta) \right\}$$ (2)

$$c_t + k_{t+1} - (1 - \delta) \bar{k}_t = A_t k_t^{\alpha(1+\eta)} n_t^{(1-\alpha)(1+\eta)}$$ (3)

$$\log A_t = \rho \log A_{t-1} + \sigma \varepsilon_t,$$ (4)

where equation (1) is the optimal labor supply condition, Euler equation (2) equates the marginal cost of reducing consumption in period $t$ and the marginal gain of consumption in the next period, equation (3) is the aggregate resource constraint in equilibrium, and equation (4) specifies the dynamics of $A_t$. The parameter $\sigma$ controls the variance of the innovation $\varepsilon_t \sim N(0,1)$ and measures the level of uncertainty in the economy. Since agents are identical, in general equilibrium the individual variables are equal to their aggregate counterparts (e.g., $k_t = \bar{k}_t$ and $n_t = \bar{n}_t$).

As shown by Benhabib and Farmer (1994), this model has a unique steady state and the steady state is a saddle if the degree of externalities is small and a sink if $\eta$ is large enough. However, Coury and Wen (2009) show that this class of models may have multiple dynamic equilibria (such as stable n-period cycles) away from the steady state even if the steady state appears to be a saddle. To be mindful of this, in the following analyses we choose the value of $\eta$ sufficiently below the critical values found by Coury and Wen (2009) for n-period cycles.

### 2.1 Second-Order Taylor Expansion

The model’s equilibrium is solved by second-order approximation methods. The variables in the above equations can be grouped into two types, the state variables and the control variables.
variables. The state variables include the capital stock and the realized exogenous shock at the beginning of each period. We denote them by the vector $s_t = (k_t, A_t)$. The control variables include consumption and hours, denoted by $z_t = (n_t, c_t)$.

Under linear-approximation methods, due to certainty equivalence, the equilibrium paths of economic variables are independent of the degree of uncertainty ($\sigma$), and the dynamic impulse responses of the model are symmetric with respect to the sign of the shocks ($A_t$). That is, the policy functions can be written as

$$z_t = g(s_t),$$

and the state variables follow the law of motion

$$s_{t+1} = h(s_t) + \zeta \sigma \varepsilon_{t+1},$$

where the $2 \times 1$ vector $\zeta = [0 \ 1]'$ since technological innovations do not directly affect the next-period capital stock.

However, under second-order approximations, second-order terms such as $E_t \hat{A}_t^2 = \frac{\sigma^2}{1-\rho^2}$ and $E_t \hat{A}_t k_{t+1}$ will emerge, so the policy functions $g(\cdot)$ and equilibrium paths of state variables $h(\cdot)$ will depend on $\sigma$. Accordingly, the general policy functions are characterized by

$$z_t = g(s_t, \sigma),$$

$$s_{t+1} = h(s_t, \sigma) + \zeta \sigma \varepsilon_{t+1}.$$

Note that when the standard deviation of the shock $\sigma = 0$, there is no uncertainty in the model and all variables remain in the non-stochastic steady state. Therefore, the impulse responses of the system to a technology shock obtained under the first-order and the second-order methods converge to each other as $\sigma$ approaches zero.

Define $\tilde{z}_t \equiv \log z_t - \log z_0$, $\tilde{s}_t \equiv \log s_t - \log s_0$. Following the literature (see, e.g., Schmitt-Grohe and Uribe, 2004), the policy functions \{g(\cdot), h(\cdot)\} can be approximated by a second-order Taylor expansion,

$$\tilde{z}_t = \left[g_s \ \ g_\sigma \right] \left[\tilde{s}_t \ \ \sigma\right] + \frac{1}{2} \left[g_{ss} \ \ 0 \ \ g_{s\sigma} \ \ 0 \ \ g_{\sigma\sigma}\right] \left[\tilde{s}_t \ \ \sigma\right],$$

$$\tilde{s}_{t+1} = \left[h_s \ \ h_\sigma \right] \left[\tilde{s}_t \ \ \sigma\right] + \frac{1}{2} \left[h_{ss} \ \ 0 \ \ h_{s\sigma} \ \ 0 \ \ h_{\sigma\sigma}\right] \left[\tilde{s}_t \ \ \sigma\right] + \left[0 \ \ 0 \ \ 0 \ \ 1\right] \sigma \varepsilon_{t+1};$$

---

8If indeterminacy occurs, one may treat period-$t$ consumption as a state variable since the steady state is a sink.
where the first derivatives $g_s$ and $g_\sigma$ are known from the solution to the first-order system. By taking second-order derivatives of the first-order conditions with respect to $s$ and $\sigma$, one can obtain a linear equation system with the unknown elements in cross-derivative coefficients $g_{ss}$, $h_{ss}$, $g_{s\sigma}$ and $h_{s\sigma}$. The cross terms are symmetric and equal to zero, $g_{s\sigma} = g_{\sigma s} = h_{s\sigma} = h_{\sigma s} = 0$, as shown by the literature. The second-order coefficient matrices are three dimensional. For example, since we have two elements in $g(s_t, \sigma)$ and two elements in $s_t$, $g_{ss}$ is thus a $2 \times 2 \times 2$ matrix.

The steady state under second-order methods is different from that under linear methods. For example, if the initial state is zero, $\hat{s}_0 = 0$, then we have $\hat{s}_1 = \frac{1}{2}h_{s\sigma}\sigma^2$; consequently the system will evolve and rest at another steady state. Moreover, at the non-stochastic steady state where $\hat{z}_t$ and $\hat{s}_t$ are zero, the policy functions differ from those under linear solutions by a constant term proportional to $\sigma^2$.

Accordingly, in order to generate impulse response functions and time series comparable to those under linear methods, one can find the stochastic steady state by numerical simulation, and then introduce shocks to generate time series or impulse responses relative to the stochastic steady-state values. Furthermore, one may note that direct use of (9) and (10) can generate an exploding path since second-order terms result in unnecessary higher-order terms in consecutive iteration. The “Pruning” process as proposed by Kim et al. (2008) is useful to overcome this problem. In fact, this process uses only first-order parts of the response to generate the second-order terms in the recursive computation.

### 2.2 A Simple Example

Suppose an economy is described by the following nonlinear equation:

$$ y_t = E_t A_{t+1} y_{t+1}^\theta, \quad |\theta| < 1; $$

(11)

where $y_t$ is an endogenous jump variable and $A_{t+1}$ an exogenous driving process satisfying

$$ \log A_t = \rho \log A_{t-1} + \sigma \varepsilon_t $$

(12)

and $\varepsilon_t \sim N(0, 1)$, where $\sigma$ denotes the standard deviation of $\varepsilon_t$. The model has two steady states: $\bar{y} = 0$ and $\bar{y} = 1$. We consider the second steady state because it is saddle-stable. To facilitate Taylor expansions around a steady state, denote $\hat{y}_t \equiv \log y_t - \log \bar{y}$ and rewrite

---

9 We depart from Schmitt-Grohe and Uribe (2004) and Kim et al. (2008) by adjusting the steady state so that the impulse responses return to zero in the long run.
equation (11) as

\[ e^{\log y_t} = E_t e^{\log A_{t+1} + \theta \log y_{t+1}}. \]  

(13)

**First-Order Method.** The first-order Taylor expansion of equation (13) around the steady state \( \{ \log \bar{y} = 0, \log \bar{A} = 0 \} \) is given by

\[ \hat{y}_t = E_t \left[ \hat{A}_{t+1} + \theta \hat{y}_{t+1} \right]. \]  

(14)

Using the method of undetermined coefficients, we can guess the solution \( \hat{y}_t = \gamma_A \hat{A}_t + \gamma_\sigma \sigma \) and substitute the solution into equation (14) to obtain

\[ \gamma_A \hat{A}_t + \gamma_\sigma \sigma = E_t \left[ \hat{A}_{t+1} + \theta \left( \gamma_A \hat{A}_{t+1} + \gamma_\sigma \sigma, \right) \right]. \]  

(15)

Applying the law of motion (12) and comparing coefficients of \( \{ \hat{A}_t, \sigma \} \) on both sides of equation (15) gives \( \gamma_A = \frac{\rho}{1-\theta \rho} \) and \( \gamma_\sigma = 0 \). So the first-order accurate solution is given by

\[ \hat{y}_t = \frac{\rho}{1-\theta \rho} \hat{A}_t. \]  

(16)

The impulse response function of \( \hat{y}_t \) to a one-standard deviation shock in \( \varepsilon_t \) can be generated from the following state-space representation of the model,

\[ \begin{pmatrix} \hat{y}_t \\ \hat{A}_t \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \rho \gamma_A \end{pmatrix} \begin{pmatrix} \hat{y}_{t-1} \\ \hat{A}_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma_A \\ 1 \end{pmatrix} \sigma \varepsilon_t, \]  

(17)

by setting \( \hat{y}_0 = 0, \hat{A}_0 = 0, \sigma = 1, \varepsilon_1 = 1 \) and \( \varepsilon_t = 0 \) for all \( t > 1 \).

**Second-Order Method.** The second-order Taylor expansion of equation (13) around the steady state \( \{ \log \bar{y} = 0, \log \bar{A} = 0 \} \) is given by

\[ \hat{y}_t + \frac{1}{2} \hat{y}_t^2 = E_t \left[ \hat{A}_{t+1} + \theta \hat{y}_{t+1} + \frac{1}{2} \left( \hat{A}_{t+1} + \theta y_{t+1} \right)^2 \right]. \]  

(18)

We can guess a second-order solution with undetermined coefficients:

\[ \hat{y}_t = \gamma_A \hat{A}_t + \gamma_{AA} \hat{A}_t^2 + \gamma_\sigma \sigma + \gamma_{\sigma \sigma} \sigma^2. \]  

(19)

Notice that substituting this second-order solution into equation (18) would generate higher-order terms such as \( \{ \hat{A}_t^x, \hat{y}_t^x, \sigma^x \} \) with \( x \geq 3 \). Since these higher-order terms are
irrelevant for our second-order solution, they can be ignored completely. That is, we can substitute out the second-order terms in equation (18) by the first-order solution in equation (16) while keeping the first-order terms as they are in equation (18).

Therefore, plugging the linear solution (16) into second-order terms in equation (18) yields

$$\hat{y}_t = E_t \left[ \hat{A}_{t+1} + \theta \hat{y}_{t+1} + \frac{1}{2} \left( \hat{A}_{t+1} + \frac{\theta \rho}{1 - \theta \rho} \hat{A}_{t+1} \right)^2 - \frac{1}{2} \left( \frac{\rho}{1 - \theta \rho} \hat{A}_t \right)^2 \right]. \quad (20)$$

Since $E_t \hat{A}_{t+1}^2 = \rho^2 \hat{A}_t^2 + \sigma^2$, the above equation reduces to

$$\hat{y}_t = \theta E_t \hat{y}_{t+1} + \rho \hat{A}_t + \frac{1}{2 (1 - \theta \rho)^2} \sigma^2. \quad (21)$$

Plugging the conjectured second-order solution (19) into equation (21) and simplifying gives

$$\gamma_A \hat{A}_t + \gamma_{AA} \hat{A}_t^2 + \gamma_\sigma \sigma + \gamma_{\sigma \sigma} \sigma^2$$

$$= \rho (1 + \theta \gamma_A) \hat{A}_t + \theta \gamma_\sigma \sigma + \theta \rho^2 \gamma_{AA} \hat{A}_t^2 + \left( \theta \gamma_{AA} + \theta \gamma_{\sigma \sigma} + \frac{1}{2 (1 - \theta \rho)^2} \right) \sigma^2. \quad (22)$$

Comparing coefficients on both sides gives \( \gamma_A = \frac{\rho}{1 - \theta \rho}, \ \gamma_\sigma = 0, \ \gamma_{AA} = 0, \text{ and } \gamma_{\sigma \sigma} = \frac{1 \rho^2}{2 (1 - \theta) (1 - \theta \rho)^2}. \) Hence, the second-order rational expectations equilibrium is given by

$$\hat{y}_t = \frac{\rho}{1 - \theta \rho} \hat{A}_t + \frac{1}{2 (1 - \theta) (1 - \theta \rho)^2} \sigma^2. \quad (23)$$

Note the second-order solution conforms to the general solution in equation (9). However, in the special model the second-order effect from $\hat{A}_t^2$ does not exist and the solution differs from the first-order solution only by a constant proportional to $\sigma^2$. Thus, the corresponding impulse response function is also the same as that of the first-order solution (up to a constant term in the steady state).

### 2.3 Calibration and Eigenvalues

We calibrate the model based on the existing literature. In particular, we set the discounting factor $\beta = 0.99$, the elasticity of intertemporal substitution $\tau = 1$, the inverse elasticity of labor supply $\gamma = 0$ (indivisible labor), capital’s share $\alpha = 0.3$ and the persistence of shock $\rho = 0.9$. 


The eigenvalues of the model is given by the linear terms and thus not affected by higher-order terms. Hence, the region of local indeterminacy is not influenced by the variance of technology shocks under second-order expansion. Excluding the exogenous driving process, the system has two eigenvalues. When externalities are small, one and only one of the eigenvalues lies inside the unit circle, thus the steady state is a saddle. When externalities are large enough so that \( \eta > \eta^* = 0.4935 \), both eigenvalues will lie inside the unit circle and the system is indeterminate (Benhabib and Farmer, 1994). Figure 1 plots the two eigenvalues as functions of the externality parameter \( \eta \). The critical value for indeterminacy is \( \eta^* = 0.4935 \).¹⁰

Figure 1. Eigenvalues.

The dashed lines correspond to the region of complex eigenvalues. It is clear from Figure 1 that the eigenvalues go through dramatic changes as the degree of externalities (\( \eta \)) increases. The explosive root jumps from positive infinity to negative infinity while the stable root goes through similar topological changes near the critical value \( \eta^* \). Both eigenvalues become stable and form a complex conjugate pair (dashed lines) for certain region of \( \eta > \eta^* \). Such dramatic changes in eigenvalues indicate that the topology of the model near the steady state is significantly altered by externalities, and that linear approximation may not be accurate enough for capturing the curvature of the equilibrium path near the steady state.

¹⁰Wen (1998a) showed that introducing capacity utilization can reduce the critical value \( \eta^* \) to about 0.1, which is more consistent with empirical estimates. Our results apply to the model of Wen (1998a), but we choose the simpler Benhabib-Farmer model for exposition purpose in this paper.
3 Dynamic Analysis

3.1 Impulse Responses

The impulse responses of the economy to a technology shock with standard deviation \( \sigma = 0.3 \) are graphed in Figure 2.\(^{11}\) For comparison, windows in the first column on the left show the impulse responses under the first-order approximation method and those in the second column on the right show those under the second-order method. The top-row windows show the results when externality are absent (\( \eta = 0 \)) and the bottom-row windows

---

\(^{11}\)If the the variance of the shock \( \sigma \) is too small, second-order and first-order methods yield very similar results. On the other hand, second-order methods may no longer be accurate if \( \sigma \) is too large. Hence, we choose a large enough \( \sigma \) so that the second-order terms are significant and at the same time the second-order solution is more accurate than linear methods. If we use the model of Wen (1998a) with variable capacity utilization, the values of \( \sigma \) can be made much smaller, in the order of \( \sigma = 0.03 \).
show the results when externalities are large enough ($\eta = 0.4$) but well below the critical value of 0.4935 for indeterminacy.\footnote{This value ($\eta = 0.4$) is also sufficiently below the threshold value for generating n-period cycles found by Coury and Wen (2009).}

It is clear from the top-row windows that first-order and second-order methods yield very similar results when externalities are not present, confirming the literature’s findings that linear solution methods provide reasonably good approximations for standard RBC models. For example, the initial impulse response of output is about 0.7 percent under first-order approximation and about 0.68 percent under second-order approximation.

However, when externalities exist, the two solution methods yield dramatically different results. The bottom-row windows in Figure 2 indicate that the initial impulse response of output is about 1.6 percent under first-order approximation but only about 0.75 percent under second-order approximation. More importantly, while the responses of output and labor remain monotonic under the first-order method, they all become hump-shaped under the second-order method, suggesting a much richer internal propagation mechanism.

When the level of externalities increase, the hump-shaped responses become even more prominent. For $\eta$ close to the critical value $\eta^*$, even a very small value of $\sigma$ can give rise to hump-shaped dynamics under the second-order method. It is thus evident that production externalities have dramatically changed the topology of the model near the steady state. It is shown by Benhabib and Wen (2004) that under linear approximation method, hump-shaped impulse responses and oscillating cycles will emerge when the externalities are large enough so that the model becomes indeterminate. Here, we show that such nonlinear dynamic may already exist when the degree of externalities are below the critical value required for indeterminacy, but they can only be captured by higher-order terms.\footnote{Coury and Wen (2009) show that the model of Benhabib and Farmer (1994) has global indeterminacy even when the model’s steady state appears to be a saddle judged by eigenvalues (i.e., for externalities below the critical level required for local indeterminacy based on first-order approximation method).}

The reason for the above results is that the coefficients of the second-order terms in a standard RBC model without externalities are in general very small (close to zero), hence linear methods usually yield quite accurate solutions. However, once externalities or increasing returns are allowed, the second-order terms can become non-negligible and very large. For example, when $\sigma = 0.3$ and $\eta = 0.4$, we have the following decision rules for capital,
consumption, and labor:

\[
\begin{bmatrix}
\hat{k}_{t+1} \\
\hat{c}_t \\
\hat{n}_t
\end{bmatrix}
= \begin{pmatrix}
0.9124 & 0.5484 \\
0.4325 & 0.9101 \\
-0.6272 & 4.4930
\end{pmatrix}
\begin{bmatrix}
\hat{k}_t \\
\hat{A}_t
\end{bmatrix}
+ \frac{1}{2}
\begin{pmatrix}
0.0463 & -0.2652 & 0.7531 \\
0.0143 & -0.1810 & 0.3924 \\
-0.7155 & 9.0494 & -19.6185
\end{pmatrix}
\begin{bmatrix}
\hat{k}_t^2 \\
\hat{k}_t \hat{A}_t \\
\hat{A}_t^2
\end{bmatrix}
+ \frac{1}{2}
\begin{pmatrix}
2.4241 \\
-0.4161 \\
20.8073
\end{pmatrix}
\sigma^2.
\]  

(24)

![Coefficient of \(\hat{A}_t^2\)](image1)

![Coefficient of \(\hat{k}_t \hat{A}_t\)](image2)

Figure 3. Second-Order Effects as \(\eta\) Increases.

Note that in labor’s decision rule the coefficient of \(\hat{A}_t^2\) is \(-19.6185\) and that of the cross term \(\hat{k}_t \hat{A}_t\) is \(9.0494\). But these coefficients are close to zero when \(\eta = 0\). For hours worked, its coefficients of \(\hat{A}_t^2\) and \(\hat{k}_t \hat{A}_t\) increase in absolute values as \(\eta\) increases towards the critical value \(\eta^*\), but both become zero again as soon as \(\eta > \eta^*\). This is shown in Figure 3 where the second-order coefficients of labor and consumption are graphed. Labor has no second-order terms under indeterminacy because hours are (log)linear function of consumption and capital, both of which are state variables under indeterminacy.

Hence, the reason that second-order method predict more subdued and hump-shaped output responses is that externalities introduce a large negative coefficient in front of the square term of technology shock, \(\hat{A}_t^2\), and a large positive term in front of the cross term between the capital stock and technology, \(\hat{k}_t \hat{A}_t\). This means that the effect of a positive technology shock on labor and output is neutralized by the square term in the impact period but enhanced by the cross term in the subsequent periods, giving rise to hump-shaped dynamic pattern.
Such a nonlinear, second-order income effect on hours also implies that the economy’s
dynamic responses to technology shocks are not symmetric. Under a negative technology
shock, the square term $\hat{A}_t^2$ reinforces the negative shock on impact while the cross term $k_t \hat{A}_t$
offset the negative effect of the shock in subsequent periods, making the impulse response
of output to a negative technology shock more contractionary but less persistent than that
predicted by the linear method (see Figure 4). For example, Figure 4 shows that the initial
drop in output is around $-1.6$ by linear approximation and about $-2.5$ by second-order
approximation. The predicted half-life is about 7 quarters under the linear method and less
than 4 quarters under the second-order method.

![Figure 4. Responses to a Negative Technology Shock.](image)

Such an asymmetric property is similar to what we observe in the U.S. economy. Ex-
panions tend to be more gradual and long-lived while contractions tend to be sharper and
short-lived (see, e.g., Neftic, 1984; and Sichel, 1993). Also, Mork (1989) and Hamilton (2003)
point out that the economy’s response to oil price shocks are highly asymmetric: a sudden
increase in oil price (adverse productivity shock) tend to depress the economy while a sud-
den decrease in oil prices tend to have little effect. Aguiar-Conraria and Wen (2007) argue
that increasing returns to scale large enough to trigger indeterminacy can explain the large
negative impact of oil price increases on the U.S. economy. Here we show that mild degree
of externalities can produce similar effect under second-order approximation.

Given the same level of $\sigma$, as $\eta$ increases, the impulse responses of output under second-
order expansion will become more and more hump-shaped and the initial value can even
become negative (which is then followed by a positive hump). This dynamic pattern provides an explanation to a puzzle existing in the indeterminacy literature: technology shocks suddenly become contractionary once the model becomes locally indeterminate.

Figure 5. Responses to a Positive Technology Shock.

The left window in Figure 5 shows the first-order approximation of the model under a positive technology shock when the steady state is indeterminate (i.e., $\eta = 0.6$),\textsuperscript{14} and the right window shows the second-order approximation of the model under a positive technology shock when $\eta = 0.48$, which is below the critical value for indeterminacy. The left window shows that technology shocks are contractionary under indeterminacy, which is puzzling because the top-left and bottom-left windows in Figure 2 indicate that externalities amplify technology shocks in the positive direction. The existing indeterminacy literature has not provided an explanation for this puzzling phenomenon.

However, the right window in Figure 5 suggests that technology shocks are already contractionary from a second-order viewpoint even before the model becomes indeterminate. Therefore, we believe that the puzzle is caused by a large income effect from increasing returns to scale on hours. When the marginal product of labor is high, it is optimal to reduce hours worked and increase leisure under the income effect but to increase hours worked under the substitution effect. However, when externalities are below the critical value $\eta^*$, the income effect is captured only by second-order terms while the substitution effect is captured by the first-order terms. Therefore, first-order method will show positive impulse responses

\textsuperscript{14}The initial consumption level is fixed at the steady state, $\hat{c}_t = 0$. See Benhabib and Wen (2004) for discussions on how to generate impulse responses to fundamental shocks in an indeterminate model.
under technology shocks. Once the model becomes indeterminate, if consumption is treated as a state variable under indeterminacy, the decision rule of labor has only first-order terms because the optimal first-order condition of labor supply (Equation 1) implies that it is a (log)linear function of the model’s state space. Hence, in this case there are no second or higher-order terms in labor and consequently, the strong income effect of a technology shock on hours can only be captured by first-order terms, explaining the puzzle in the left window of Figure 5.

3.2 Behavior of Investment

Based on linear methods, it has become well-known in the RBC literature that investment is extremely volatile under technology shocks (see, e.g., Kydland and Prescott, 1982). However, this is not necessarily the case from a second-order perspective (see Figure 6).

Figure 6. Responses of Investment to a Positive Shock.
Because of the asymmetric second-order income effect, investment is very volatile only under a negative technology shock but not so under a positive technology shock. In fact, if the elasticity of labor supply is large enough or technology is not highly persistent, the initial response of savings (investment) to a positive shock can be very mild or even negative, as shown in Figure 6. Notice that in each window of Figure 6, second-order approximation always yields a lower level of investment than first-order approximation. This is due to an extremely large negative coefficient of the square term $\hat{A}_t^2$ in the investment decision rule, $\left[ \hat{i}_t \right] = \left( -2.5029 \ 21.9379 \right) \left[ \hat{k}_t \ \hat{A}_t \right] + \frac{1}{2} \left( -10.1104 \ 139.2464 \ -439.1138 \right) \left[ \hat{k}_t^2 \ \hat{k}_t \hat{A}_t \ \hat{A}_t^2 \right] + \frac{96.9648}{2} \sigma^2$. By the same token, however, investment will always appear to be more volatile than that predicted by linear method under a negative technology shock.

### 3.3 Second-Order Effects of Sunspots

Since the coefficients of second-order terms are obtained based on the linear solution (see Schmitt-Grohe and Uribe, 2004; and Lombardo and Sutherland, 2007), the zone of indeterminacy will remain unchanged if indeterminacy arises in the model. Farmer and Guo (1994) simulate the Benhabib-Farmer (1994) model under indeterminacy by linear method and show that sunspot shocks can generate business-cycle comovements among output, consumption, investment, and hours. The effects of positive and negative sunspot shocks under linear method are symmetric and monotonic (see the left-column windows in Figure 7, where the upper row pertains to positive shock and the lower row pertains to negative shock).

However, as shown in the right-column windows, under second-order approximation, positive and negative sunspot shocks have asymmetric effects on the economy. In particular, under a negative sunspot shock (the lower right-corner window), the economy has hump-shaped impulse responses, while under a positive sunspot shock (the upper right-corner window), the economy has no initial hump but tends to overshoot its steady state from above. The hump-shaped impulse responses of output to negative sunspot shocks are in contrast to the analysis of Schmitt-Grohe (2000), where she argues under linear solution method that sunspots shocks cannot generate hump-shaped output dynamics and forecastable comovements among output, consumption, hours and investment. Here we show that this may not necessarily be the case if the second-order effects are taken into consideration.

---

15 When recursive equilibria fail to exist, the perturbation methods employed by the literature and this paper cannot be applied (see Peralta-Alva and Santos, 2009, for the issues involved). Here we follow Benhabib and Farmer (1994) and the existing literature by taking as given the existence of recursive equilibria in the region of indeterminacy.
Notice in Figure 7 both second-order and first-order methods yield exactly the same initial magnitude of responses for all variables under a positive sunspot shock. The results differ only in the subsequent periods. This is so because with indeterminacy consumption is a state variable and the coefficients of second-order terms for labor are zero; thus, in the first period the responses of consumption and labor are only due to first-order effect while the capital stock stays unchanged. Consequently, sunspot has only a first-order effect on the responses of output in the impact period. For example, the decision rules under sunspot
shocks take the following form when \( \sigma_s = 0.3 \) and \( \eta = 0.6 \),

\[
\begin{bmatrix}
\hat{k}_{t+1} \\
\hat{c}_{t+1} \\
\hat{n}_{t+1}
\end{bmatrix} = 
\begin{pmatrix}
0.5070 & 1.0000 \\
-0.1304 & 1.2229 \\
-4.0000 & 8.3333
\end{pmatrix}
\begin{bmatrix}
\hat{k}_t \\
\hat{c}_t \\
\hat{n}_t
\end{bmatrix} + \frac{1}{2}
\begin{pmatrix}
2.5900 & -9.7502 & 9.1002 \\
0.0327 & 0.1184 & -0.2984
\end{pmatrix}
\begin{bmatrix}
\hat{k}_t^2 \\
\hat{k}_t \hat{c}_t \\
\hat{c}_t^2
\end{bmatrix} + \frac{1}{2}
\begin{pmatrix}
0 & 8.5717 \\
8.5717 & 0
\end{pmatrix} \sigma_s^2 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sigma \varepsilon_{st+1},
\]

(25)

where \( \sigma_s^2 \) is the variance of \( \varepsilon_{st} \equiv \hat{c}_t - E_{t-1} \hat{c}_t \), the one-period ahead forecasting error of consumption (sunspots). Clearly, starting from the steady state where \( \hat{k}_t = \hat{k}_{t-1} = \hat{c}_{t-1} = 0 \), the effects of a sunspot shock on consumption in period \( t \) is determined only by the forecasting error \( \varepsilon_{st} \) and not by any higher-order terms in the state space. On the other hand, the coefficients of the second-order terms are zero in the decision rule of hours. In addition, output is a linear function of capital and labor. Therefore, both second-order method and first-order method yield the same initial impulse responses for consumption, hours and output. However, in the subsequent periods the responses under the two methods diverge significantly. In particular, because the cross term \( \hat{k}_t \hat{c}_t \) has a large negative effect and the square term \( \hat{c}_t^2 \) has a large positive effect on the capital stock, the impulse responses of output and labor become hump-shaped downward under a negative sunspot shock but, under a positive sunspot shock, they become monotonically decreasing and over-shooting the steady state from above.

4 Accuracy Test

As pointed out by Jin and Judd (2002), second-order approximation is not necessarily better than linear methods in terms of solution accuracy, depending on the models used and the parameter regions. Judd (1998) proposes to use Euler Equation Error (EEE) as a criterion for non-local accuracy test, which is expressed as the logarithm of the Euler equation residual

\[
EEE = \log_{10} \left| 1 - \left\{ \beta E_t \left[ c_{t+1}^{-\tau} \left( \alpha A_{t+1} k_{t+1}^{\alpha (1+\eta)} n_{t+1}^{(1-\alpha)(1+\eta)} - 1 + \delta \right) \right] \right\}^{\frac{1}{\tau}} \right|,
\]

(26)

where \( \{ c_{t+1}, c_t, n_{t+1}, k_{t+1} \} \) are determined by the second-order policy rules discussed above. Note that EEE is in general negative, and a smaller (i.e., more negative) EEE implies an
improved accuracy. For example, EEE is negative infinity (−∞) when the solution is exact (100% accurate).

Figure 8 plots the EEEs of the model under the first-order and the second-order method, respectively, when σ = 0.3 and η = 0.4. The vertical axes represents EEE, the right-front axes represents deviations of 𝑃, from its steady state 0 in both positive and negative directions, and the left-front axes represents the deviations of capital from its steady state 𝑘 = 0. Since the first-order EEE lies everywhere above the second-order EEE both at the steady state and when the model is significantly away from the steady state, the second-order solution dominates the linear method in terms of accuracy under the current calibrations.
On the other hand, if we fix the technology level at its steady state ($\hat{A}_t = 0$) or the capital stock at its steady state ($\hat{k}_t = 0$), and let the degree of externalities vary, the top-row windows in Figure 9 shows that the degree of accuracy deteriorates as $\eta$ increases toward $\eta^* = 0.4935$. In particular, the second-order method is not necessarily better than the linear method around the critical point of indeterminacy, $\eta^*$. However, for most values of $\eta$, the second-order method dominates the first-order method. The bottom window in Figure 9 is a two-dimensional picture of the top-row windows at the point where $\hat{k}_t = 0$ and $\hat{A}_t = 0$. It shows that the accuracy of both the first-order and the second-order approximation methods deteriorates as the externality $\eta$ increases toward the critical point of indeterminacy, but with
the second-order method more accurate than the linear method except around the point of indeterminacy.

5 Conclusion

This paper shows that externalities and IRS can have important implications for business-cycle dynamics that have not been fully appreciated by first-order approximation methods. In particular, they induce a strong second-order asymmetric income effect on leisure so that hump-shaped output dynamics may emerge even if externalities are sufficiently below the critical value required for local indeterminacy. Similarly, i.i.d sunspots shocks under indeterminacy may be able to generate forecastable comovements of output, hours, consumption and investment, in contrast to the conclusion reached by Schmitt-Grohe (2000). Moreover, this asymmetric second-order income effect is able to generate business cycle dynamics that are qualitatively consistent with the asymmetric nature of the business cycle observed in the U.S. economy. Whether similar results can also be found in models with other types of frictions, such as borrowing constraints, remain to be investigated.
References


