

A Defense of RBC: Understanding the Puzzling Effects of Technology Shocks

Pengfei Wang
Hong Kong University of Science & Technology
Clear Water Bay, Hong Kong

Yi Wen
Federal Reserve Bank of St. Louis
St. Louis, MO 63146 USA

First Version: April 2007

This Version: July 2007

Abstract

The research led by Gali (AER 1999) and Basu et al. (AER 2006) raises two important questions regarding the validity of the RBC theory: (i) How important are technology shocks in explaining the business cycle? (ii) Do impulse responses to technology shocks found in the data reject the assumption of flexible prices? Using an RBC model, this paper argues that the conditional impulse responses of the U.S. economy to technology shocks are not grounds to reject the notion that technology shocks are the main driving force of the business cycle and the assumption of flexible prices, in contrast with the conclusions reached by the literature. Our paper also provides a new approach to deriving aggregate production functions.

Keywords: RBC, Technology Shocks, Aggregate Production Function, Entry and Exit, Sticky Prices, Demand Rigidity, Business Cycle.

JEL codes: E31, E32, E52.

*We thank Luke Shimek and George Fortier for research assistance. The usual disclaimer applies. Correspondence: Yi Wen, Research Department, Federal Reserve Bank of St. Louis, St. Louis, MO, 63144. Phone: 314-444-8559. Fax: 314-444-8731. Email: yi.wen@stls.frb.org.

1 Introduction

Because of intertemporal substitutions and instantaneous market clearing, standard RBC models imply a sharp rise in aggregate labor and investment, as well as the real interest rate, immediately after an aggregate technology shock. However, Gali (1999) and Basu et al. (2006) found that aggregate technology shocks in the U.S. economy are contractionary to labor, investment, and the real interest rate in the short run. This finding has two important implications regarding the validity of the RBC theory: (i) aggregate technology shocks may not be the main driving force of the business cycle because aggregate labor and investment are procyclical in the data; (ii) aggregate supply is not responsive to technology shocks in the short run, suggesting sticky prices. These implications have led this literature to conclude that RBC theory is dead (see, e.g., Francis and Ramey, 2005).¹

It is possible that technology shocks are not important and prices are sticky. However, based on the above findings alone one cannot logically conclude that this is indeed the case in reality. Whether or not technology shocks are an important driving force of the business cycle does not follow logically from the sign of the initial impulse responses to technology shocks. Two time series can still be positively correlated at the business cycle frequency even if they have the opposite signs of impulse responses on impact.²

The importance of technology shocks notwithstanding,³ a contractionary effect of technology shocks on aggregate inputs and factor prices does not reject flexible prices. This point is the main focus of the paper.

Based on the "purified" technology series estimated by Basu et al. (2006), we confirm that both aggregate technology shocks and sector-specific technology shocks are contractionary on sectorial activities. However, sectorial inputs decrease only temporarily under aggregate technology shocks but permanently under sector-specific technology shocks. In other words, while aggregate technological progress is contractionary in the short run, sector-specific technological progress tends

¹There is a fast growing literature regarding the dynamic effects of technology shocks. Important works include Basu (1998), Basu et al. (2002, 2006), Chang and Hong (2006), Chari et al. (2005), Christiano et al. (2004, 2006), Fernald (2007), Fisher (2002, 2006), Francis and Ramey (2005), Gali (1999), Gali et al. (2003), Gali and Rabanal (2004), Liu and Phaneuf (2006), Ramey (2004), Vigfusson (2004), and Uhlig (2004), among many others. The bulk of this literature supports the empirical finding that technology shocks have a negative impact on employment and hours worked in the short run. However, Christiano et al. (2004) argue that hours' response to technology shocks is positive if they enter the VAR in levels, instead of in growth rate. But this finding is disputed by Gali and Rabanal (2004) and Fernald (2007). In particular, Fernald (2007) shows that, after allowing for (statistically and economically significant) trend breaks in productivity, hours fall when technology improves, regardless how hours enter the VAR.

²For example, two sine curves with a phase difference may still comove together despite opposite signs at the origin.

³See Fisher (2006) for an alternative approach and argument regarding the importance of investment-specific technology shocks, and see the discussion of Gali and Rabanal (2004) regarding Fisher's work.

to be contractionary in both the short and long run. We show that these stylized facts are fully consistent with a flexible-price RBC with firm entry and exit.

Our RBC model with firm entry and exit is motivated by the fact that aggregate net business formation is strongly procyclical under aggregate technology shocks, suggesting the number of firms is an important margin for aggregate output and inputs to adjust under business cycle shocks. Since this margin of capacity adjustment lies only at the aggregate level but not at the firm level, aggregate and firm-level technology shocks should have asymmetric implications for resource allocation and economic fluctuations. We illustrate this intuition using a perfectly competitive flexible-price RBC model in which the number of firms is a key propagation mechanism of aggregate technology shocks.

Francis and Ramey (2005) show that a flexible-price RBC model with aggregate demand rigidity (namely, habit formation and investment adjustment costs) can also generate short-run negative labor responses to technology shocks. They also show that assuming a Leontief aggregate production technology (with labor and capital as perfect complements) can achieve similar results. But these models with aggregate rigidities are not able to generate short-run negative investment responses to technology shocks, which is one of the key stylized facts of the U.S. economy emphasized by Basu et al. (2006).

Our model does not suffer from this shortcoming. In addition, the dynamics of the real wage and the real interest rate under technology shocks are emphasized by Francis and Ramey (2005), Basu et al. (2006), and Liu and Phanuef (2007) as important litmus tests for business cycle models. In the data, a positive aggregate technology shock leads to a modest rise on impact and a permanent rise in the long run for the real wage, and a sharp decrease in the real interest rate in the short run. These stylized facts are viewed by this literature as consistent with sticky prices/wages, but not with flexible prices/wages. However, our model is consistent with the dynamic behavior of the real wage and the real interest rate despite the lack of price-wage stickiness in the model. We view this as an advantage of our model because sticky price models imply a systematic positive relationship between the degree of price stickiness and the extent of the contractionary effects of technology shocks on hours. Empirical evidence at the industry level for the existence of such a relationship is absent (see Chang and Hong, 2006).

Our approach draws inspiration from the existing literature by emphasizing rigidity in factor demand. However, we build demand rigidity into the micro level without assuming demand rigidities at the aggregate level. A micro-level rigidity in factor demand can arise from a Leontief production structure at the firm level due to fixed capacities. Such a micro structure is consistent with standard aggregate production technologies with positive elasticity of substitution across aggregate inputs (see, e.g., Houthakker 1955-56, Levhari 1968, Johansen 1972, and Lagos 2006). Since the number of firms can vary because of entry and exit, our model with micro-level rigidity is identical to a standard

frictionless RBC model in aggregate dynamics, everything else equal. In this case, permanent aggregate technology shocks generate immediate positive responses of aggregate investment and labor, which is inconsistent with the data. However, if firms must wait for one period after entry to produce (or to earn profits) because of time-to-build, then the model starts to behave very differently from standard frictionless RBC models and is able to explain all of the aforementioned facts about the puzzling effects of aggregate and sector-specific technology shocks. Hence, time-to-build is the only assumption we need in order to break the equivalence of aggregate dynamics between our model and a standard frictionless RBC model, despite demand rigidities at the micro level.

An additional advantage of our model is that many types of standard aggregate production functions, such as the Dixit-Stiglitz function with constant returns to scale and constant elasticity of substitution (CES) across aggregate inputs, can be derived as special cases from our model. Hence, the assumption of a Leontief structure at the firm level is innocuous. In light of this, our approach provides a micro foundation for standard RBC models that assume CES or Cobb-Douglas aggregate production technologies.⁴

The rest of the paper proceeds as follows. Section 2 presents stylized facts about the dynamic effects of technology shocks on aggregate and sectorial activities. These stylized facts appear to be profoundly inconsistent with RBC models with flexible prices. However, in Sections 3 through 5, we show this is not the case: RBC models with perfect competition and instantaneous market clearing are fully consistent with the stylized facts. A simple one-sector benchmark model is presented in Section 3 to gain intuition, and a full model with multiple sectors and heterogenous firms is presented in Sections 4 and 5. Section 6 shows that our model with firm entry and exit based on a Leontief structure of firms provides a general-equilibrium micro foundation for standard aggregate production functions with constant returns to scale and constant elasticity of substitution across aggregate inputs. Our approach also provides an alternative explanation for the total factor productivity (TFP), in contrast with the labor search approach of Lagos (2006). Section 7 concludes.

2 Stylized Facts

In this section, we replicate three sets of stylized facts (Basu et al. 2006) regarding the dynamic effects of technology shocks. First, in the initial period after an aggregate technology shock, aggregate consumption rises, aggregate output and net business formation do not change significantly, while aggregate investment and labor decline sharply. In the longer run, however, all variables

⁴The strategy of relying on a micro-level Leontief structure to derive standard aggregate production functions is also used by Lagos (2006) to study the micro foundations of aggregate TFP. However, since Lagos (2006) uses a labor search model, the conditions for deriving aggregate production functions in his approach are very different from ours.

increase permanently. Second, sectorial output does not change significantly while sectorial inputs decrease significantly in the initial period after an aggregate technology shock; but they all rise permanently in the longer run. Third, under sector-specific technology shocks, sectorial output does not change significantly (in either the short or long run), but sectorial inputs decrease in both the short and long run. These stylized facts are robust to various methods of estimation and identification. They suggest that aggregate technology shocks are contractionary in the short run but expansionary in the long run at both aggregate and sectorial levels, but sector-specific technology shocks are contractionary for sectorial activity in both the short run and the long run.

Estimation Method. All variables are defined in growth rates. Time series representing "technology shocks" are taken from Basu et al. Three alternative methods are used to identify innovations in the technology series and generate impulse responses from the U.S. economy. The first approach follows Basu et al. by assuming that the estimated technology series are completely exogenous. Namely, we estimate a restricted bivariate VAR for the technology series () and another variable under interest ():

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t-1 \\ t-1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} t-2 \\ t-2 \end{pmatrix} + \begin{pmatrix} t \\ t \end{pmatrix} \quad (1)$$

where f and g are OLS coefficients and v_t is the innovation to technology. This restricted VAR implies that v_t explains one hundred percent of the variance in the technology series (). Notice that v_t and ε_t are orthogonal by construction.⁵ The second and third approaches do not assume one hundred percent exogeneity of the estimated technology series and are based on an unrestricted VAR:

$$\begin{pmatrix} t \\ t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t-1 \\ t-1 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} t-2 \\ t-2 \end{pmatrix} + \begin{pmatrix} x_t \\ y_t \end{pmatrix} \quad (2)$$

where x_t and y_t are OLS residuals with covariance matrix Σ . To identify innovations to technology, we use both the Cholesky decomposition and the Blanchard-Quah (1989) method, respectively, to construct the mapping:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 11 & 12 \\ 21 & 22 \end{pmatrix} \begin{pmatrix} t \\ t \end{pmatrix} \quad (3)$$

where v_t is the identified innovation to technology (orthogonal to the innovation ε_t). The variance of both innovations is normalized to one. Thus, the above mapping implies $\Sigma = \Gamma\Gamma'$. Under the Cholesky decomposition, Γ is triangular with $12 = 0$. Under the Blanchard-Quah method, in the moving average representation, $v_t = \sum_{j=0}^{\infty} \alpha_j v_{t-j} + \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}$, the long-run restriction

⁵The second projection in Equation (1) implies v_t is orthogonal to $\{x_t, x_{t-1}, x_{t-2}\}$. Hence, it must also be orthogonal to $\varepsilon_t = x_t - a_1x_{t-1} - a_2x_{t-2}$.

$\sum_{j=0}^{\infty} \alpha_j = 0$ is imposed so that only α_0 can have a permanent effect on y_t . The results are very similar under the three different approaches, which help to establish the robustness of the stylized facts we try to document.

Data. All sectorial data are from Basu et al. (2006), which also clarifies the construction of the aggregate technology series as the weighted average of sectorial technology series. Aggregate output, consumption, and investment are taken from the Bureau of Economic Analysis. Aggregate employment and hours are taken from the Bureau of Labor Statistics. Net business formation is taken from the Department of Commerce (Bureau of the Census) and Dun & Bradstreet, Inc. To be consistent with technology series from Basu et al., all data are annual averages and are truncated to match the sample period of 1949-1996.⁶

2.1 Impulse responses to aggregate technology shocks

Figure 1 displays the impulse responses of aggregate output, net business formation, consumption, investment, and hours to an aggregate technology shock. The shaded area in each window represents one-standard-deviation bands.

⁶Output is defined as real GDP, consumption as total real consumption, investment as total non-residential fixed investment. All data are measured as year-over-year percentage change (Table 1.1.1 from BEA). Labor statistics include total nonfarm employees and total private average weekly hours of production workers. Both data series are monthly. In forming the year-over-year growth rate, we use the monthly average of each year, consistent with Basu et al. (2006). Since data on hours are not available before 1964, we have also included total nonfarm employees in Figure 1 (panel 5). The real wage is the ratio of the nominal hourly wage and the consumer price index (CPI). The real interest rate is the difference between the federal funds rate and the CPI inflation rate.

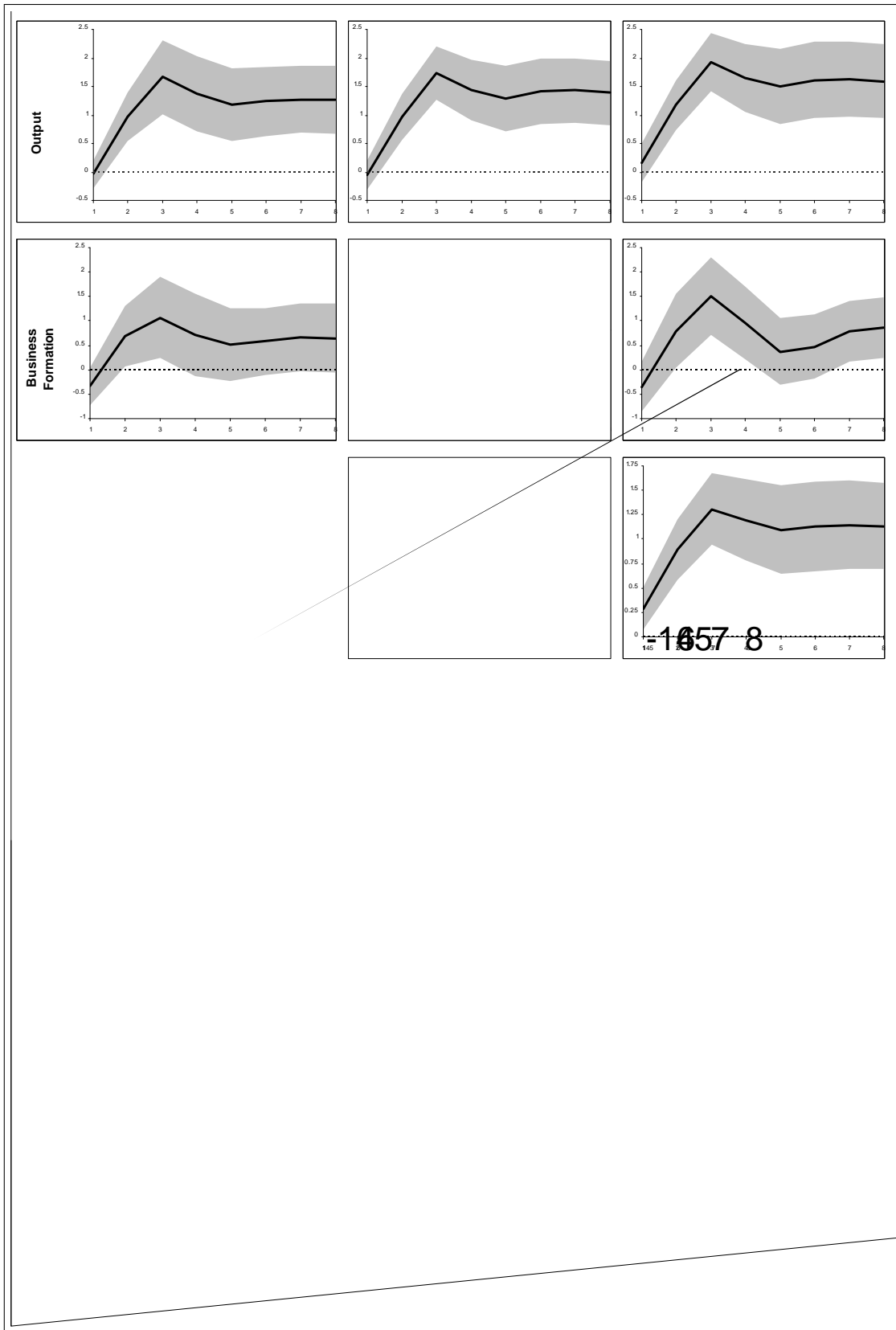


Figure 1. Impulse Responses to Aggregate Technology Shock.

The first column shows estimations based on the restricted VAR of Basu et al. The second and third columns show estimations based on the Cholesky and Blanchard-Quah methods, respectively. After a technology shock, GDP does not respond significantly on impact but increases permanently afterwards. Net business formation declines slightly on impact but increases (comoves with output) in the future. Consumption rises on impact and rises further subsequently toward its long-run steady state. Investment and employment both decline sharply on impact, but then rise significantly in the longer run. Similarly, hours decline sharply on impact but increase strongly with a lag. These dynamic patterns of impulse responses are consistent with the findings of Basu et al. (2006).⁷

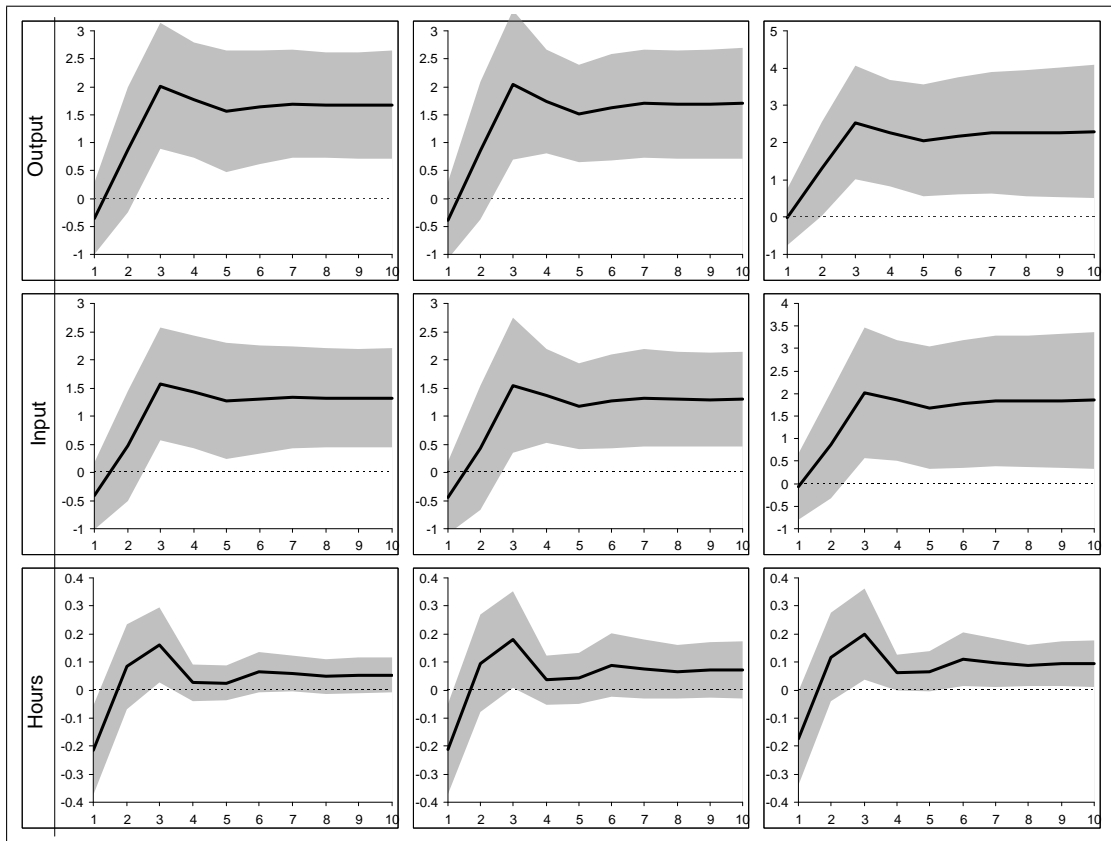


Figure 2. Sectorial Responses to Aggregate Technology Shock.

Figure 2 shows the impulse responses of sectorial output and sectorial inputs to aggregate technology shocks. As before, the three columns correspond to the three different methods of shock identification. In each window, the solid line represents the average of all 29 sectors' impulse responses and the shaded area represents one-standard-deviation bands across sectors. The first panel pertains to output, the third panel to hours, and the middle panel to other inputs (e.g.,

⁷When the TFP is purified using the method of Basu et al., Chang and Hong (2006) obtain similar results with respect to the impact of technology shocks on hours.

capital and employment).⁸ The figure shows that the effects of aggregate technology shocks on sectorial output and inputs broadly mimic those on aggregate output and inputs. Namely, in the initial period after the shock, sectorial output does not change significantly but sectorial inputs (especially hours) tend to decline sharply. However, output and inputs all increase permanently in the longer run.

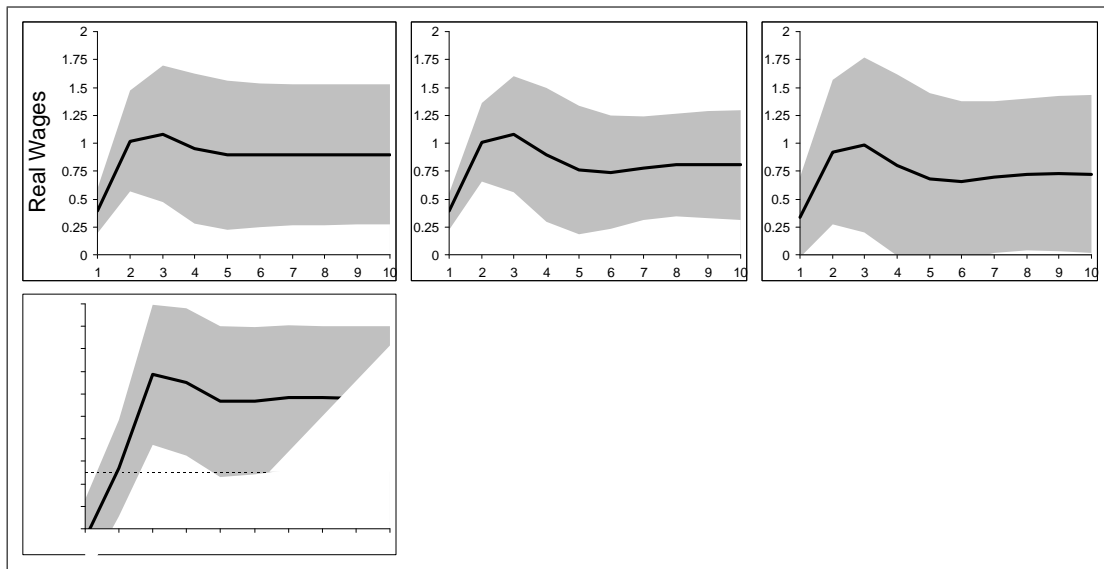


Figure 3. Impulse Responses of Real Wage and Real Rate.

Figure 3 shows the impulse responses of the real wage and the real interest rate to technology shocks. The most striking difference between the two series is that the real wage increases while the real rate decreases on impact after a technology shock. Since the real factor price is proportional to the output-factor ratio times the real marginal cost and since output (as well as the capital stock) does not respond significantly in the initial period after the shock, the short-run behaviors of the real wage and the real interest rate imply that the real marginal cost must decrease under technology shocks but no more than the decrease in hours. This imposes a severe constraint and discipline on a business cycle model. Standard RBC models with constant marginal cost are inconsistent with such dynamic behaviors of factor prices.⁹ Also notice the dramatic difference in the magnitude of the responses of the two factor prices: the real interest rate is about 10 times more volatile than the real wage. This is obviously consistent with models where labor is more responsive than capital to technology shocks, provided that one important condition is met: the negative response of marginal cost to technology shocks must be highly transitory. In the initial period, labor decreases but the capital stock does not move; hence, the decline in the marginal cost brings down the real interest

⁸In the data of Basu et al., there are two types of inputs: i) hours worked and ii) all else (including intermediate goods, capital and labor, etc.).

⁹For example, a model with habit formation and investment adjustment cost implies a positive response of the real interest rate on impact.

rate more than the real wage. If changes in the marginal cost are highly transitory, then as output increases in the intermediate run, the closer comovement of labor to output compared with capital to output will cause the real wage to increase less than the real interest rate. In other words, the real interest rate would not increase more than the real wage in the longer run if the initial decrease in marginal cost were highly persistent under permanent technology shocks.

2.2 Impulse responses to sector-specific technology shocks

Measure of sector-specific technology shocks. Ideally, a sector-specific technology shock process should be orthogonal not only to aggregate technology shocks but also to the sector-specific technology shocks in other sectors of the economy. However, the measures of sectorial technology of Basu et al. do not satisfy these criteria for two reasons. First, the aggregate technology series of Basu et al. is constructed as a weighted average of the sectorial technology series; hence, it is correlated with all sectorial technology series by construction. Second, the sectorial technology series are not orthogonal among each other. The sample period of the data (48 observations in each series) is short and the number of sectors (29 sectors) is relatively large; therefore, constructing sector-specific technology shocks as residuals of each sectorial technology series regressing on the other 28 sectorial series and on the aggregate series as dependent variables is troublesome because of potential colinearity problems. In this paper, we purify the sectorial technology series of Basu et al. as much as we can by regressing each sectorial series on its own lag and the aggregate technology series.¹⁰ Although this purification procedure does not necessarily yield orthogonality among the "purified" sector-specific technology shock processes, it does improve the degree of purity by ensuring that all sectorial technology shocks are orthogonal to the aggregate technology. However, since the constructed "sector-specific" technology shocks are not necessarily orthogonal among each possible pair of sectors, the degree of purity can be judged only by the average correlations among these series.

Figure 4 shows the distribution of correlations between any two series of purified sector-specific technology shocks. Since there are 29 sectors in our sample, there are 406 possible pair-wise combinations. The histogram of the 406 correlations in Figure 3 shows that the constructed sector-specific technology shock series are not pair-wise orthogonal. The maximum correlation is 0.8 and the minimum correlation is -0.6. However, the distribution is approximately normal, with the mass centering around zero (the mean of the correlations is 0.04). Given this, although the technology shock series are not one hundred percent "sector-specific", they have zero correlations on average across sectors. Hence, regarding the dynamic effects of sector-specific technology shocks on sectorial activities, the average impulse responses across all sectors maybe more informative than the individual impulse responses.

¹⁰The results are very similar when the lagged variable is excluded.

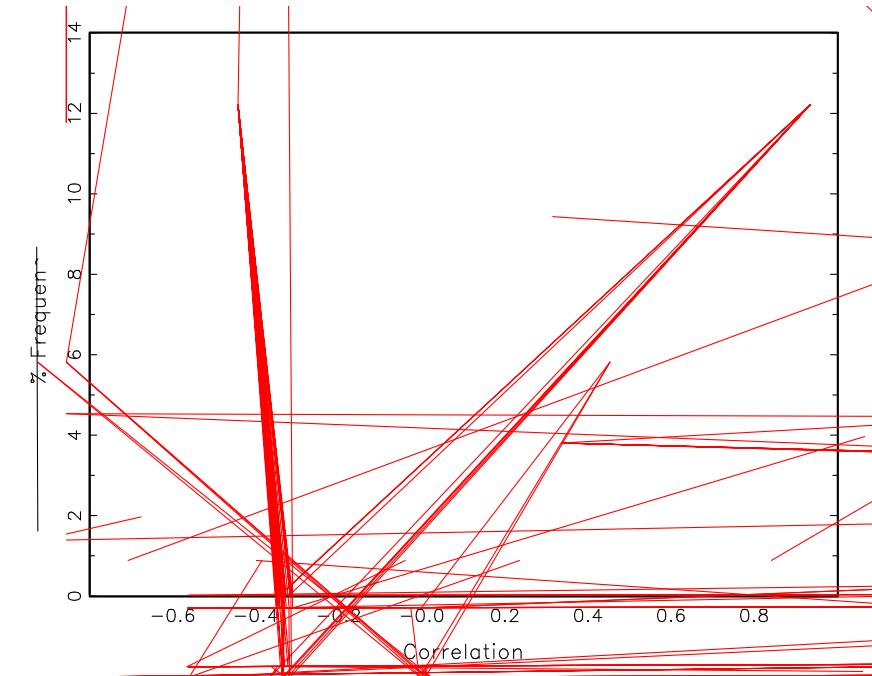


Figure 4. Distribution of Correlations among Sector-Specific Technology Shocks.

Figure 5 shows the impulse responses of sectorial output and inputs to sector-specific technology shocks of the same sector, where solid lines represent the average impulse response and the shaded area represents one standard deviation from the mean across sectors. As before, in each panel the first column corresponds to the method of Basu et al. and the second and third columns correspond to Cholesky decomposition and the method of Blanchard and Quah, respectively. The first panel shows that, on average, sector-specific technology shocks have little impact on sectorial output in both the short run and the long run (especially under the method of Basu et al.). This is in sharp contrast to the effects of aggregate technology shocks shown in Figure 2. The second and the third panel show that sector-specific technology shocks are contractionary to sectorial inputs in both the short and the long run, which is also in sharp contrast to the effect of aggregate technology shocks shown in Figure 2 above.

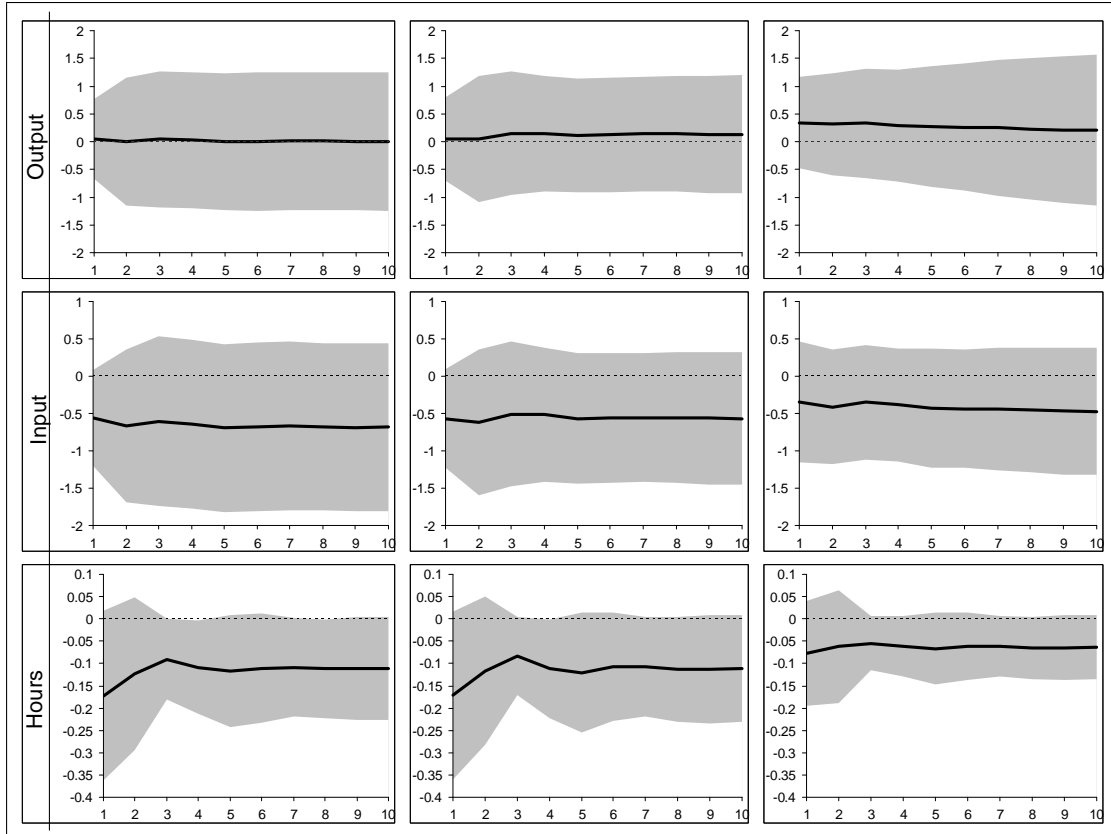


Figure 5. Sectorial Responses to Sector-Specific Technology Shocks.

Intuitively, the permanent contractionary effects of sector-specific technology shocks on sectorial inputs make sense because it would be in firms' best interests to permanently reduce expenditures on inputs when output could not be changed. Hence, the fundamental question is why sectorial output remains essentially unchanged over time under sector-specific technology shocks, but rises sharply in the longer run under aggregate technology shocks?

Our approach to answer this question is to model real rigidities at the firm level but allow for full flexibility at the aggregate level. The aggregate flexibility is achieved by allowing for firm entry and exit. Our model with firm entry and exit is identical to a standard frictionless one-sector RBC model in aggregate dynamics if there is no time-to-build. In this case, aggregate technology shocks generate positive initial impulse responses for both output and inputs, which is inconsistent with the data as shown by Gali (1999) and Basu et al. (2006). However, if it takes time (say one period) for new firms to set up production plants and earn profits after entry, then our model starts to behave very differently from the standard RBC model and is able to explain the empirical facts regarding the effects of technology shocks at both the sectorial and aggregate levels.

3 Benchmark Model

3.1 Final Good

The economy produces only one type of final good (). There are many identical final good producers in any period , with each producing only a fixed quantity of the final good. Without loss of generality, this quantity is normalized to one. In a sense, each firm can be viewed as a production assembly line with fixed production capacity. There is a fixed cost, $\Phi \in (0, 1)$, to enter the final good industry. Entry and exit under perfect competition will determine the total mass (number) of final good producers, Ω_t , in each period. The intermediate good for producing is . Producing one unit of the final good requires units of , where is a constant. Without loss of generality we can normalize both and the price of the final good () to one, $w = p_y = 1$. Hence, the production function is simply $y = x$ and the aggregate output simply equals the total number of producers. Let w_x be the price of input. A final good producer's profit maximization problem is

$$\max_x (x - w_x x) \tag{4}$$

This yields the demand for input:

$$= \begin{cases} 1 & \text{if } w_x \leq 1 \\ 0 & \text{if } w_x > 1 \end{cases} \tag{5}$$

Profit in each period for each producer is given by

$$= \begin{cases} 1 - w_x & \text{if } w_x \leq 1 \\ 0 & \text{if } w_x > 1 \end{cases} \tag{6}$$

In each period the aggregate supply of output () is determined by the number of firms (production lines) and is equal to $\int_0^\Omega = \Omega$, and the aggregate demand for input is

is illustrated in Figure 6.

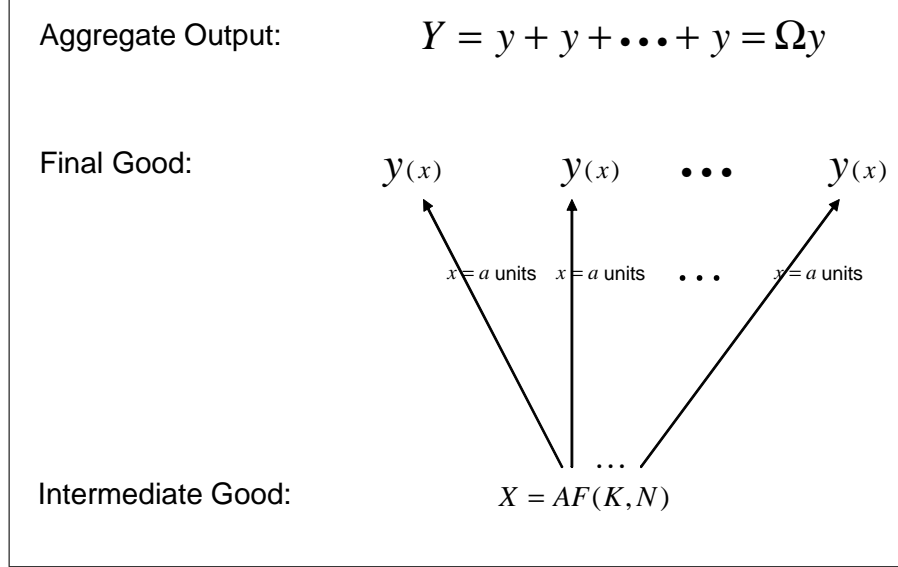


Figure 6. Production Structure.

Firm entry and exit. The setup of firms' entry and exit is similar to that in Ghironi and Melitz (2005). In each period, there are potentially infinite entrants, which make the final good industry perfectly competitive. The one-time fixed entry cost (Φ) is paid in terms of the final good. After entry, each firm faces a stochastic probability of exit, $\theta \in (0, 1)$. The probability of exit is assumed to depend on the size of aggregate technology shocks. We assume that firms must wait one period to produce output after entry because of time-to-build. The value of a firm in period t is then determined by

$$v_t = \theta v_{t+1} + (1 - \theta) \sum_{j=1}^{\infty} \theta^{j-1} \left[\prod_{i=1}^j (1 - \theta_{t+i}) \right] \Lambda_{t+j+1} v_{t+j+1} \quad (7)$$

where Λ_{t+j} is the ratio of marginal utilities between period $t+j$ and period t . We can also write this equation recursively as

$$v_t = \theta v_{t+1} + (1 - \theta) v_{t+1} \quad (8)$$

Free entry then implies $v_t = \Phi$. The evolution of the number of final good producers is

$$\Omega_{t+1} = (1 - \theta) \Omega_t + \theta \quad (9)$$

where θ is the number of new entrants in period t .¹¹

¹¹The fixed cost (Φ) can be interpreted, for example, either as a regulation fee or as the cost of purchasing structural capital goods such as buildings or production lines. In the latter case, the probability of exit (θ) can be interpreted as the depreciation rate of the structural capital.

3.2 Intermediate good

In the benchmark model there is only one type of intermediate good, x . The intermediate good market is also perfectly competitive. For simplicity, we assume there are no costs to enter this market. The production function of a representative producer of the intermediate good is

$$x_t = A_t z_t^\alpha k_t^{1-\alpha} n_t \quad (10)$$

where A_t stands for TFP (technology) shocks, and k_t and n_t stand for capital and labor. The unit cost of labor is the real wage, w_t , and the unit cost of renting capital from households is the sum of the real interest rate and the depreciation rate, $r_t + \delta$. Profit maximization by a representative firm gives $\frac{x_t}{k_t} = r_t + \delta$ and $(1 - \alpha) \frac{x_t}{n_t} = w_t$. Perfect competition implies that price equals the marginal cost, $p_t = \frac{1}{A_t} \left(\frac{r_t + \delta}{\alpha} \right)^\alpha \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha}$. Since the intermediate good sector has only one representative firm, the aggregate supply of intermediate goods is x_t .

3.3 Household

A representative household receives interest income by renting capital to intermediate good producers and wage income by working. It also receives net profit income (gross profits minus fixed entry costs),

$$\Pi_t = \int_{i=0}^{\Omega} p_t x_{i,t} - \Phi \quad (11)$$

from final good producers, where Ω is the number of existing incumbents and Φ is the number of new entrants. The household's utility maximization problem is standard:

$$\max \sum_{t=0}^{\infty} \beta^t [\log(c_t) + \log(1 - c_t)] \quad (12)$$

subject to

$$k_{t+1} = k_t + (1 - \delta)k_t - k_t + \Pi_t \quad (13)$$

The first-order conditions for the household are $\beta^t c_t^{-1} = (\beta^{t+1} c_{t+1})^{-1}$ and $\beta^t k_t^{-1} = \beta^{t+1} k_{t+1}^{-1} (1 + r_{t+1})$

3.4 General equilibrium

A general equilibrium is a set of quantities and prices, $\{k_t, n_t, x_t, p_t, c_t, \Pi_t\}_{t=0}^{\infty}$ such that firms maximize profits, households maximize utilities, and all markets clear. The resource constraint (15) for the representative household in equilibrium becomes

$$t + t_{t+1} (1 - \delta) t + t\Phi = t \quad (14)$$

where the aggregate output is determined by the number of final good producers in equilibrium, $t = \Omega_t$, which in turn is also the total demand for intermediate goods, $\Omega_t = t = t \frac{a}{t} t^{1-\alpha}$. The system of equations that determine the general equilibrium thus consists of the aggregate production function, $t = t \frac{a}{t} t^{1-\alpha}$, the resource constraint (14), and

$$\Phi = t \frac{t}{t+1} (t_{t+1} + (1 - \delta) \Phi) \quad (15)$$

$$\Omega_{t+1} = (1 - \delta) \Omega_t + t \quad (16)$$

$$t = 1 - x_t \quad (17)$$

$$x_t \frac{t}{t} = t + \quad (18)$$

$$(1 - \delta) x_t \frac{t}{t} = t \quad (19)$$

$$t^{-1} = (1 - \delta)^{-1} \quad (20)$$

$$t^{-1} = t \frac{-1}{t+1} (1 + t_{t+1}) \quad (21)$$

3.5 Equivalence to standard RBC model

Suppose final good producers can start production and earn profits within the same period of entry and the probability of exit after production is one ($\delta = 1$). Then equation (8) becomes $t = t$. Hence equations (15)-(17) collapse to $t = \Phi$, $x_t = 1 - \Phi$ and $t = \Omega_t$. The aggregate resource constraint (14) becomes

$$t + t_{t+1} (1 - \delta) t = (1 - \Phi) t \frac{a}{t} t^{1-\alpha} \quad (22)$$

and the factor prices become $t + = (1 - \Phi) \frac{Y_t}{K_t}$ and $t = (1 - \delta) (1 - \Phi) \frac{Y_t}{N_t}$. Since the marginal cost $x_t = 1 - \Phi$ (0 < 1) is constant, equations (18)-(22) indicate that the dynamics of this model are the same as those implied by a standard frictionless RBC model (e.g., King, Plosser, and Rebelo, 1988).

Hence, time-to-build is the only assumption we need to break the equivalence in aggregate dynamics between our model and a standard RBC model. In the real world, it takes time to set up production facilities. According to Kydland and Prescott (1982), the average construction period

for plants is about two years. In addition to the construction period, there are other forms of cost of entry. For example, Djankov et al. (2002) present data on the regulation of entry of start-up firms in 85 countries. The data cover the number of procedures, official time, and official cost a start-up must bear before it can legally operate. They found that the official costs of entry are extremely high in most countries, ranging from under 0.5 percent of per capita GDP in the United States to over 4.6 times per capita GDP in the Dominican Republic, with the worldwide average of 47 percent of annual per capita income. They also found that the minimum time needed between application (paying the entry fees) and proper operation of the business after entry is about one and one-half months on average across countries, assuming there are no delays by either the applicant or the regulators.

3.6 Impulse responses

Calibration. The time period is a year. Let $\beta = 0.96$, $\delta = 0.4$, $\alpha = 0.1$ and the steady-state fraction of hours worked $\bar{h} = 0.2$ (or about 35 hours per week). Let the fixed cost of entry $\Phi = 0.1$ (which implies the steady-state share of entry cost to GDP is $0.1(1 - \alpha)$). The results are not sensitive to these parameter values. To ensure stationarity of n_t , assume the probability of exit depends on the innovation (rather than the level) of technology shocks, $\log(\delta_t) = -\eta \log(\epsilon_t)$. The probability of exit corresponds to the business failure rate in the real world. Based on data from the U.S. (1949-96), one percent increase in the aggregate technology reduces the business failure rate by about six percent in the long run; hence, we set $\eta = 6$. This negative elasticity implies that a positive aggregate technology shock reduces the probability of exit because of improved efficiency for all firms. To calibrate the steady-state value of \bar{m} , we note that the dynamics of the model variables, except the first-period response of the number of new entrants (n_t), are not sensitive to this parameter. For example, the initial responses of investment and hours are negative for any value of $\bar{m} > 2$ ($0 < \bar{m} < 2$). However, the initial impulse response of n_t is sensitive to this parameter: it is positive when \bar{m} is small enough but negative when \bar{m} is large enough. The average business failure rate (at annual frequency) for the U.S. economy is about 710 per 10,000 listed enterprises, implying a steady-state value close to $\bar{m} = 0.1$.¹² Under this value, the initial impulse response of n_t is positive. However, for larger values of \bar{m} , the model can generate a negative initial response of n_t to technology shocks. The long-run responses of n_t are always positive regardless of \bar{m} . For this reason, we simulate the model using two alternative values, $\bar{m} = 0.1$ and 0.25 . These values imply a steady-state markup in the range of 1.5–3%, which is well within the empirical estimates suggested by the literature.

¹²This is also the value adopted by Ghironi and Melitz (2005).

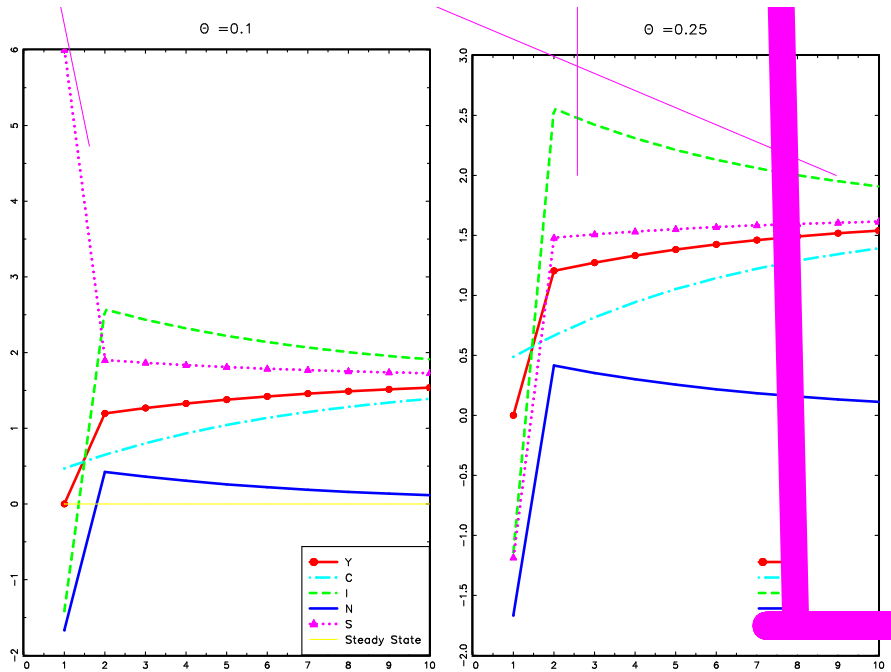


Figure 7. Impulse Responses to a Permanent Technology Shock.

Figure 7 shows that the aggregate dynamics of the benchmark model are broadly consistent with the U.S. data. The left window in Figure 7 shows that, in the initial period after a permanent increase in aggregate technology, aggregate output does not change, net business formation increases sharply, consumption rises gradually, and investment and hours fall sharply. In the longer run, however, all variables increase permanently to a higher steady state. Output does not respond to technology shocks in the initial period because it is determined by the exiting number of firms when a shock hits the economy, which is a state variable determined by net business formation in the past. This short-run rigidity of output is caused by time-to-build; namely, new firms cannot produce output immediately after entry. Because of this short-run rigidity of aggregate output, demand for aggregate labor must decrease, as in a sticky price model. Consumption rises immediately after the shock because consumers correctly anticipate the increase in permanent income. Given that current income (output) is fixed, aggregate savings (investment) must fall to support the increase in consumption. Since a positive technology shock increases expected profits by reducing the marginal cost, net business formation tends to rise immediately after the shock.

The right window in Figure 7 shows that with a higher value of $\bar{\theta}$ the dynamics of net business formation change dramatically while the dynamics of the other variables remain essentially unchanged. Under a higher value of $\bar{\theta}$ net business formation may decrease in the initial period

After a positive technology shock, the planner opts to increase consumption as a result of a higher permanent income. Given that the current income (output) level is fixed, whether it is optimal to increase or decrease the investment in structural capital (i.e., the number of new firms) depends on the cost and the benefit. The benefit is the increased production capacity in the future, which implies higher future output. But adding new firms is also costly (the entry cost), which decreases current consumption. When $\bar{\theta}$ is large, the benefit is reduced because of a higher depreciation rate of firms. In addition, technology shocks increase the survival rate of firms in the short run ($\theta > 0$), hence mitigating the needs for investment in new firms. Consequently, the initial investment in new firms decreases. For values of $\bar{\theta}$ in between, net business formation does not respond significantly to technology shocks on impact.¹³

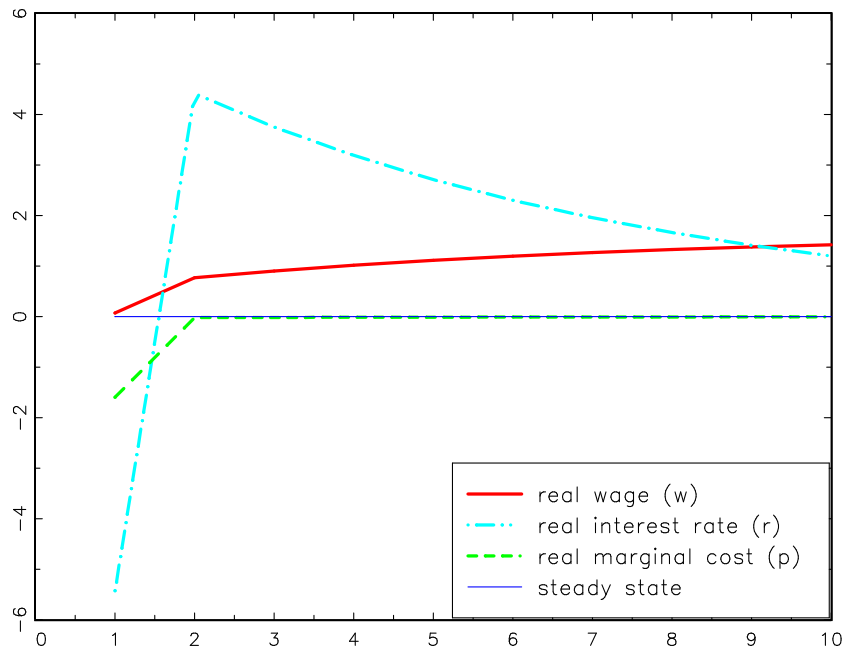


Figure 8. Impulse Responses of Factor Prices.

The impulse responses of the real wage, real interest rate, and real marginal cost are graphed in Figure 8.¹⁴ It shows that a positive technology shock leads to a moderate increase in the real wage and a sharp decrease in the real interest rate on impact, consistent with the data. Also consistent with the data is the prediction that the real interest rate is about 10 times more volatile than the real wage in the intermediate run. The initial drop in the real interest rate is due to the decline in the marginal cost. Notice the highly transitory decline in marginal cost. This feature enables the

¹³The intuition that changing the value of $\bar{\theta}$ has little effect on the impulse responses of the economy, except that of net business formation, can be understood from the fact that s_t is a flow variable while Ω_t is a stock variable and the fact that most aggregate variables are closely related to Ω_t rather than s_t .

¹⁴The dynamics of factor prices are not sensitive to the value of $\bar{\theta}$.

real interest rate to be more volatile than the real wage in the short run because hours are more capable than capital of comoving with output over time. Entry and exit are the key to generating a time-varying marginal cost in our model.

4 Multisector Model

To explain the dynamic effects of technology shocks on sectorial activity, in this section we extend the benchmark model to a multisector economy with exactly the same type of micro-rigidity. Assume that producing one unit of the final good requires a continuum of different types of intermediate goods with measure one, $\int_{j=0}^1 g_j^1$. The production function is

$$= \int_{j=0}^1 j \quad (23)$$

The price of j is denoted p_j . As in the benchmark model, the demand for j is rigid:

$$j = \begin{cases} j & \text{if } j \leq 1 \\ 0 & \text{if } j > 1 \end{cases} \quad (24)$$

where j is a constant. Hence, the vector $\{j\}$ can be viewed as the input-output coefficients of the firm. The gross profit function for a final good producer is $\pi = \int_0^1 p_j j$.

Except for expanding the input type from one to many, the structure of the model is similar to the one-sector benchmark model. For example, the value of the firm is still $v_t = \lambda_{t+1} (v_{t+1} + (1 - \lambda_{t+1}) v_{t+1})$ and the law of motion for the number of final good firms is still $\Omega_{t+1} = (1 - \lambda_t) \Omega_t + \lambda_t$. As before, each firm in the final good sector can produce only a fixed quantity of output. Without loss of generality, this quantity is also normalized to one as in the benchmark model, $\int_{j=0}^1 j = 1$. Since by assumption the production capacity of each firm is fixed, to produce more output, more firms are needed. Thus the aggregate supply of output is still determined by the number of firms:

$$y_t = \int_{i=0}^{\Omega_t} 1 = \Omega_t.$$

The production functions for intermediate goods are similar across sectors:

$$g_j = g_j (z_j, z_j) \quad (25)$$

where z_j represents a sector-specific technology shock process orthogonal to other sectorial technology shocks z_i ($i \neq j$) and the aggregate technology shock process z . The structure of the economy is illustrated in Figure 9.

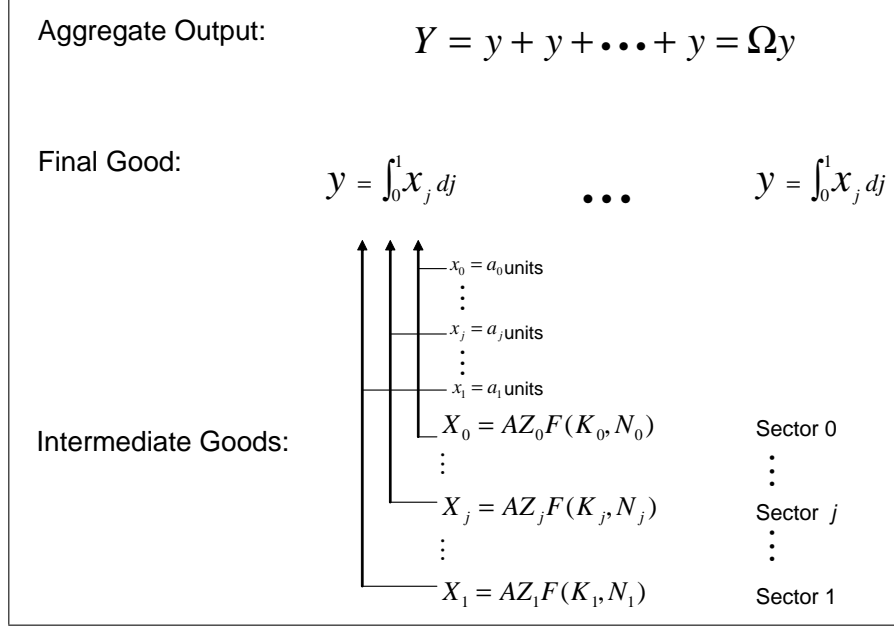


Figure 9. Multisector Model.

As in the benchmark model, a representative household receives interest payments by renting capital to intermediate producers and wages from working. It also receives net profit income ($\Pi = \int_{i=0}^{\Omega} \Phi$) from all final good producers, where Ω is the number of existing incumbent firms and Φ is the number of new entrants. The household's utility function is identical to the benchmark model and the budget constraint is given by $t + t_{+1} = t + (1 + r) t + \Pi_t$ where $t = \int_0^1 p_j$ and $t_{+1} = \int_0^1 p_{j+1}$. The first-order conditions for the household are the same as before. Notice the first-order conditions are identical across j .

Profit maximization for each intermediate-good firm in sector j gives the following first-order conditions, $\frac{X_j}{K_j} = r + \delta$ and $(1 - \alpha) \frac{X_j}{N_j} = w$. Combining with the production function gives the marginal cost of good j :

$$p_j = \frac{1}{j} \left(\frac{w}{r + \delta} \right)^\alpha \left(\frac{1}{1 - \alpha} \right)^{1-\alpha} \quad (26)$$

The aggregate output can also be expressed as

$$Y = \int_{i=0}^{\Omega} \left(\int_{j=0}^1 p_j \right) = \int_{j=0}^1 (p_j \Omega) \quad (27)$$

where $p_j \Omega$ is the aggregate demand for intermediate good j . The aggregate supply of intermediate good j is X_j . Hence, the ratio between any two types of intermediate goods is constant, $\frac{X_j}{X_i} = \frac{a_j}{a_i}$

The first-order conditions for labor and capital then imply $\frac{Z_j K_j}{a_j} = \frac{Z_i K_i}{a_i}$ and $\frac{Z_j N_j}{a_j} = \frac{Z_i N_i}{a_i}$. Rearranging and integrating over j on both sides of the equations gives $y_j = \left(\int_0^1 \frac{a_i}{Z_i} \right) \frac{a_j}{Z_j}$ and $n_j = \left(\int_0^1 \frac{a_i}{Z_i} \right) \frac{a_j}{Z_j}$. Note that the aggregator $\left(\int_0^1 \frac{a_i}{Z_i} \right)$ is independent of j for all j given the orthogonality assumption among sectorial shocks. The distribution of y_j can be chosen such that the normalization, $\left(\int_0^1 \frac{a_i}{Z_i} \right) = 1$, holds. Thus we have

$$y_j = \frac{y}{j} \quad (28)$$

$$n_j = \frac{n}{j} \quad (29)$$

Substituting y_j and n_j into the sectorial production function, $y_j = j^\alpha \left(\frac{n}{j} \right)^{1-\alpha}$, gives

$$y_j = j^{\alpha - 1 - \alpha} \quad (30)$$

In equilibrium the final good production function becomes¹⁵

$$Y = \int_{j=0}^1 (y_j \Omega) = \int_{j=0}^1 y_j = \int_{j=0}^1 j^{\alpha - 1 - \alpha} \quad (31)$$

4.1 Impulse responses

Impulse responses of aggregate variables, such as output, consumption, investment, and hours, to aggregate technology shocks are exactly the same as in the benchmark one-sector model. Impulse responses of sectorial output and inputs to both aggregate and sector-specific technology shocks can be inferred from equations (28)-(30). First, sectorial output (y_j) and inputs (n_j) are proportional to aggregate output (y) and aggregate inputs (n), respectively. Hence, the impulse responses of sectorial output and sectorial inputs behave similarly to aggregate output and inputs, respectively, under aggregate technology shocks. This is consistent with the data. Second, sectorial technology Z_j affects only sectorial inputs but not sectorial output. Hence, under sector-specific technology shocks, sectorial output remains constant but sectorial inputs decrease when Z_j increases, as in the data.

¹⁵ Alternatively, we can derive the aggregate final good production function from the market clearing condition, $\alpha_j \Omega = X_j$ for all j . This implies $\Omega (= Y) = \frac{X_j}{a_j} = AK^\alpha N^{1-\alpha}$.

4.2 Equivalence to standard RBC model

Denote $\pi_t = \frac{1}{A} \left(\frac{r+\delta}{\alpha}\right)^\alpha \left(\frac{w}{1-\alpha}\right)^{1-\alpha}$. The gross profit of each final good producer is then $\pi_t = 1 - \int_0^1 p_{jt} j_t = 1$. Utilizing the expressions for $p_{jt} = j_t$ and j_t , we can show that the market prices for capital and labor can be written as $r_t + \delta = \frac{Y_t}{K_t}$ and $w_t = (1 - \alpha) \frac{Y_t}{N_t}$. Given that there are a total of n_t new entrants in period t , the aggregate net profits from all final good producers are then $(1 - \alpha) n_t \Phi$. The household's resource constraint becomes identical to (14), $c_t + i_{t+1} = (1 - \delta) i_t + n_t \Phi$.

Suppose firms can start production immediately upon entry and $\pi_t = 1$. These assumptions imply $\pi_t = 1 - \Phi$ and $\Omega_t (= n_t) = n_t$. The aggregate resource constraint then becomes identical to (22), $c_t + i_{t+1} = (1 - \delta) i_t = (1 - \Phi) n_t$. Clearly, the dynamics of aggregate output, consumption, investment, and hours under aggregate technology shocks in this model are now equivalent to a standard frictionless one-sector RBC model.

5 Explaining Heterogeneity

Although our model is broadly consistent with the stylized facts reported in Section 2, it lacks the ability to explain the heterogeneous responses across sectors. Namely, after a positive sector-specific technology shock, factor demand decreases only on average, but in some sectors the factor demand increases (see, e.g., Figure 5). In this section, we show that the model can be easily extended to account for this heterogeneity of dynamic responses of sectorial activity to technology shocks.

As before, assume that producing a final good requires a continuum of different types of intermediate goods, $\int_0^1 g_{jt}^1$. However, the input-output coefficient in equation (24), p_{jt} , is now re-interpreted as the productivity of one unit of intermediate good j in producing the final good. Under this interpretation, p_{jt} can be considered as a random draw from a distribution $p_j(\cdot)$ in sector j . Hence, the production function of a final good firm i can be rewritten as

$$y_i = \int_0^1 p_j(\cdot) g_{jt}^1 \quad (32)$$

where the index function $p_j(\cdot)$ reflects the rigidity in factor demand: $p_j(\cdot) = 1$ if $\cdot = j$ and $p_j(\cdot) = 0$ if $\cdot \neq j$. Namely, each final good firm in each period draws an idiosyncratic random productivity for each type of intermediate good, so that firm i can transform one unit of intermediate good j into $p_j(\cdot)$ units of final good with probability $p_j(\cdot)$. Notice that we assume the distribution function differs across sector: $p_j(\cdot) \neq p_k(\cdot)$ if $j \neq k$. Denote $p_j(p_j) = 1 - \int_{a \geq p_j} p_j(\cdot) p_j(\cdot)$ as the probability that $p_j(\cdot) \geq p_j$. Since the demand for each type of intermediate good is either

zero or one, the aggregate demand for intermediate good j , by the law of large numbers, is then given by

$$X_j = \int_0^{\Omega} (x_j) = [1 - F_j(\omega_j)] \Omega \quad (33)$$

The negative of the price elasticity of demand for j is determined by

$$\eta_j = \frac{j}{1 - F_j(\omega_j)} > 0 \quad (34)$$

Impulse responses to sector-specific technology shocks. As before, the demand functions for labor and capital by each firm in the intermediate good sector j are determined, respectively, by the first-order conditions, $\omega_j L_j = (1 + \tau) w_j$ and $(1 - F_j) K_j = r_j$. The price of intermediate good j is still given by $p_j = \frac{1}{AZ_j} \left(\frac{r+\delta}{\alpha}\right)^\alpha \left(\frac{w}{1-\alpha}\right)^{1-\alpha}$. Since sector-specific technology shocks (ω_j) do not affect the real wage and the real interest rate, the changes of factor demand around the steady state for capital and labor are given, respectively, by

$$\hat{K}_j = (\eta_j - 1) \hat{\omega}_j \quad (35)$$

$$\hat{L}_j = (\eta_j - 1) \hat{\omega}_j \quad (36)$$

where a circumflex denotes log-linearization around the steady state. Clearly, under a sector-specific technology shock, demand for capital and labor will decrease in sector j if $\eta_j > 1$, and will increase in sector j if $\eta_j < 1$. In particular, the changes are permanent if the shocks are permanent. These predictions are consistent with the stylized facts presented in Figure 5. Notice that the model collapses to the previous one if $\eta_j = 0$ for all j . The average impulse response across sectors is determined by $\bar{\eta}_j - 1 = \int_0^1 \eta_j - 1$. The data suggest that $\bar{\eta}_j > 1$, which is the assumption we make.

Hence, allowing for heterogeneity in the distribution of (ω_j) can explain the heterogeneous responses of inputs across sectors. The question is: Will this affect the pattern of impulse responses of the model to aggregate technology shocks? This question is addressed below.

Impulse responses to aggregate technology shocks. Using equation (33), the ratio between any two types of intermediate good demand is $\frac{X_j}{X_i} = \frac{1-F_j}{1-F_i}$. The first-order conditions for capital and labor then imply $\frac{Z_j K_j}{1-F_j} = \frac{Z_i K_i}{1-F_i}$ and $\frac{Z_j N_j}{1-F_j} = \frac{Z_i N_i}{1-F_i}$. Rearranging and integrating over $\omega \in [0, 1]$ on both sides of the two identities gives the relationships, $\bar{K}_j = \left(\int_0^1 \frac{1-F_i}{Z_i}\right)^{-1} \frac{1-F_j}{Z_j}$ and $\bar{N}_j =$

$\left(\int_0^1 \frac{1-F_i}{Z_i}\right)^{-1} \frac{1-F_j}{Z_j}$. In the absence of sector-specific technology shocks ($Z_j = 1$ for all j), we have $Z_j = 1$ for all j and

$$Z_j = (1 - \alpha_j) \quad (37)$$

$$Z_j = (1 - \alpha_j) \quad (38)$$

where $\left(\int_0^1 (1 - \alpha_j)\right)^{-1}$. Substituting these relationships into the sectorial production function (with $Z_j = 1$), $Z_j = \alpha_j^{1-\alpha}$, gives

$$Z_j = (1 - \alpha_j)^{\alpha - 1 - \alpha} \quad (39)$$

Comparing with equation (33) gives the equilibrium number of firms,

$$\Omega = \alpha^{1-\alpha} \quad (40)$$

The factor demand functions for capital and labor can also be rewritten as

$$\frac{\alpha^{1-\alpha}}{\Omega} = \quad (41)$$

$$(1 - \alpha) \frac{\alpha^{1-\alpha}}{\Omega} = \quad (42)$$

Clearly, the aggregate dynamics of factor prices are similar to those in the previous section.

By changing the order of integration and by the law of large numbers, the aggregate output is given by

$$= \Omega \int_0^1 j(\cdot) \quad (43)$$

where $j(\cdot) = \int_{a \geq p} j(\cdot)$ is the expected (average) marginal product of intermediate good and $\int j$ is the average final output per firm. Similarly, the gross profit for the final good sector is

$$\begin{aligned} \Pi &= \int_0^\Omega \left(\int_0^1 \left(\left(\cdot \right) \left(\cdot \right) \right) \right) \\ &= \Omega \int_0^1 [j \quad (1 - \alpha_j)] \end{aligned} \quad (44)$$

Hence, the average (expected) profit of a final good firm is determined by

$$= \int_0^1 [j - (1 - j)] \quad (45)$$

The value of a firm is given by $v_t = \frac{\lambda_{t+1}}{\lambda_t} (v_{t+1} + (1 - \lambda_{t+1}) v_{t+1})$. Free entry implies $v_t = \Phi$. The household's problem is the same as before (see the Appendix for details of log-linearizing the model and parameter calibration).

Figure 10 shows that, for the average price elasticity of demand less than one (e.g., $\int_0^1 \epsilon_j = 0.2$), the impulse responses of aggregate variables to an aggregate technology shock mimic those of the previous model discussed in Section 4, with the exception that the initial response of output is not exactly zero but positive.¹⁶ The model is able to explain several key stylized facts of the business cycle emphasized by Kydland and Prescott (1982): (i) Aggregate consumption, investment, and hours comove with output; and (ii) aggregate consumption is less volatile while aggregate investment is more volatile than output. For example, in terms of annual growth rate, the ratio of standard deviation of consumption to output is 0.61 in the model and 0.62 in the data; the ratio of standard deviation of investment to output is 3.2 in the model and 3.1 in the data. The match could not be better. The correlations of consumption, investment, and hours to output are 0.55, 0.88, and 0.61, respectively, in the model, and 0.75, 0.80, and 0.42 for their counterparts in the U.S. economy. The match is quite good, although not perfect. This brings us back to the first point raised in the beginning of the paper regarding the importance of technology shocks in explaining the business cycle: Namely, the fact that investment and hours fall sharply on impact under technology shocks does not logically imply that they are countercyclical with respect to output. Things also depend on whether their forecastable future movements (in the terminology of Rotemberg and Woodford, 1996) move together with that of output. Figure 10 shows that investment and hours are expected to move with output in the future beyond the impact period, despite the fact they have opposite signs for initial responses with respect to output. Hence, investment growth and hours growth are procyclical with respect to output growth, as in the data.

¹⁶Since aggregate output is determined by $Y = \Omega \int G_j dj$, even if the number of firms (Ω) does not change initially, technology shocks can affect aggregate output by changing individual firms' expected payoff $\int_{a \geq p} af(a) da$ when the marginal cost (p) changes. The effect is larger the larger the price elasticity of firms' demand. Hence, when the average price elasticity of demand approaches infinity, the model starts to behave like a standard RBC model; namely, labor, investment, and the real interest rate increase sharply on impact. Notice that an infinitely large price elasticity of demand implies that prices are very sticky relative to demand. Hence the implication of our model is the opposite of the sticky price model.

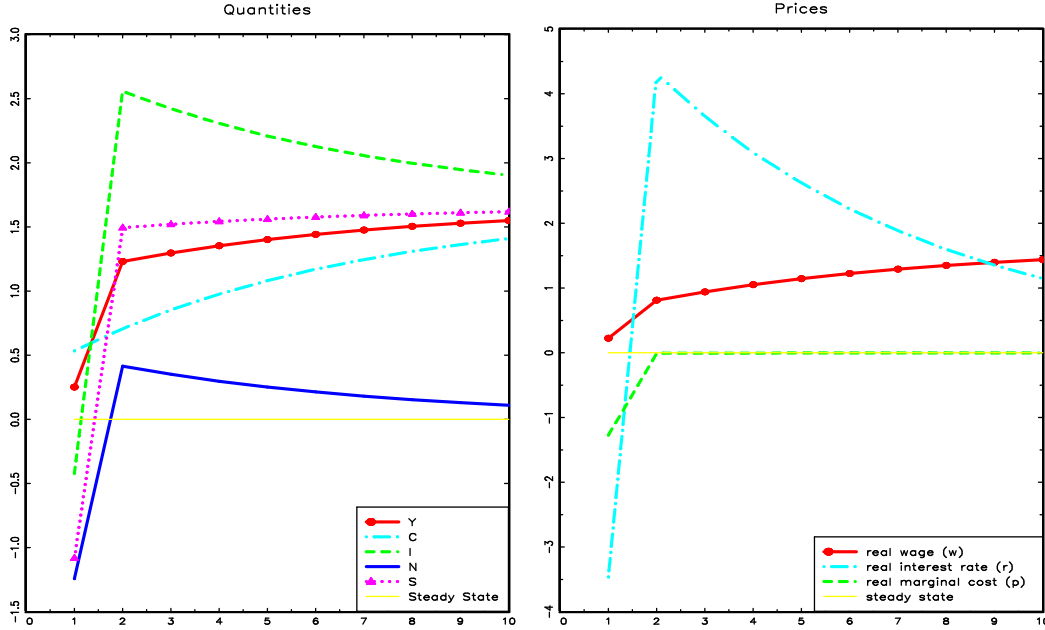


Figure 10. Impulse Responses to an Aggregate Technology Shock.

6 Aggregate Production Functions

A micro-level rigidity in factor-demand does not by itself imply any aggregate rigidities, as long as the number of firms can respond to aggregate shocks on impact. Hence, the aggregate dynamics of our model are identical to those of a frictionless RBC model if there is no time-to-build component. To further illustrate the usefulness and macroeconomic implications of our model, this section shows that many familiar aggregate production functions with constant returns to scale and positive elasticity of substitution across aggregate inputs can be derived as special cases from our model with firm entry and exit, despite the factor-demand rigidities at the firm level.

6.1 The CES Function

Consider the production function in (32), $y_i = \int_0^1 (\cdot) (\cdot)$. Assume that (\cdot) follows the Pareto distribution $(\cdot) = 1 - (\frac{a_{\min}}{a})^\sigma$ for all \cdot with the location parameter $a_{\min} > 0$ the shape parameter $\sigma > 0$ and the support $\cdot \in [a_{\min}, 1)$. Without loss of generality, assume $a_{\min} = 1$. Hence, $(\cdot) = 1 - (\frac{1}{a})^\sigma$. It follows immediately that the aggregate demand for intermediate good is given by

$$j = \left(\frac{1}{j}\right)^\sigma \Omega \quad (46)$$

which has a constant price elasticity of σ . The aggregate output in this economy is given by $Y = \int_0^{\Omega} \left(\int_0^1 (c_j)^{\frac{\sigma-1}{\sigma}} (p_j)^{\frac{1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$. By changing the order of integration and by the law of large numbers, we have $Y = \Omega \int_0^1 c_j^{\frac{\sigma-1}{\sigma}}$ where $c_j = \int_{a_{i,j} \geq p_j} (c_j)^{\frac{\sigma-1}{\sigma}} (p_j)^{\frac{1}{\sigma}}$ is the expected output of using intermediate good c_j for each firm. Under the Pareto distribution, we have $c_j = \frac{\sigma}{\sigma-1} \left(\frac{1}{p_j} \right)^{\sigma-1}$. Hence, the aggregate output is

$$Y = \frac{\Omega}{1} \int_0^1 \left(\frac{1}{j} \right)^{\sigma-1} = \frac{\Omega}{1} \int_0^1 \frac{j^{\frac{\sigma-1}{\sigma}}}{j^{\frac{\sigma-1}{\sigma}}} \quad (47)$$

where the second equality is obtained under (46). If $\sigma > 1$ then c_j is well defined.

The expected profit of a final good firm is given by

$$\pi = \int_0^1 [c_j - j(1 - (c_j)^{\frac{\sigma-1}{\sigma}})] = \frac{1}{1} \int_0^1 \left(\frac{1}{j} \right)^{\sigma-1} = \frac{1}{\Omega} \quad (48)$$

Assuming no time-to-build and the probability of exit after production is one ($\delta = 1$), free entry then implies $\pi = \Phi$. Hence, we have $\Omega = \frac{Y}{\sigma\Phi}$. Substituting this into the aggregate output in (47) and re-arranging gives the Dixit-Stiglitz CES aggregate production function,

$$Y = (\Phi) \left(\int_0^1 \frac{j^{\frac{\sigma-1}{\sigma}}}{j^{\frac{\sigma-1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}} \quad (49)$$

where $(\Phi) = \left(\frac{\sigma}{\sigma-1} \right)^{\frac{\sigma}{\sigma-1}} \left(\frac{1}{\sigma\Phi} \right)^{\frac{1}{\sigma-1}}$. Notice that there is an additional constant, (Φ) , in front of the production function. This constant is a function of the parameter σ and depends negatively on the fixed cost (Φ), demonstrating that the conventional measure of TFP can depend on market structures of the economy. In particular, trade policies affecting the cost of entering a market can affect TFP. This result is in line with those of Lagos (2006).

A special case of the above example is when the production function of the final good firm is given by

$$y_i = k_i + l_i \quad (50)$$

where k_i is capital, l_i is labor, and f_i, g_i are independent random draws from a common distribution. The factor demand for capital and labor are given, respectively, by

$$k_i = \begin{cases} f_i & \text{if } g_i < 1 \\ 0 & \text{otherwise} \end{cases} \quad (51)$$

$$l_i = \begin{cases} g_i & \text{if } g_i < 1 \\ 0 & \text{otherwise} \end{cases} \quad (52)$$

where r and w stand for prices of capital and labor, respectively. The output of a particular firm is then

$$y_i = r \alpha_i (k_i) + w (1 - \alpha_i) (l_i) \quad (53)$$

where (k_i) and (l_i) are index functions. The expected profit of a final good firm is determined by

$$\pi_i = \int_r^\infty (k_i) (l_i) + \int_w^\infty (k_i) (l_i) \quad (54)$$

With the Pareto distribution, it follows immediately that the aggregate demand for capital and labor are determined, respectively, by

$$k = \Omega \left(\frac{1}{r} \right)^\sigma \quad (55)$$

$$l = \Omega \left(\frac{1}{w} \right)^\sigma \quad (56)$$

The aggregate output is then given by

$$\begin{aligned} Y &= \left[\int_r^\infty (k_i) + \int_w^\infty (k_i) \right] \Omega \\ &= \frac{1}{\sigma-1} \left[\left(\frac{1}{r} \right)^{\sigma-1} + \left(\frac{1}{w} \right)^{\sigma-1} \right] \Omega \end{aligned} \quad (57)$$

Since the expected profit is $\pi = \frac{1}{\sigma-1} \left[\left(\frac{1}{r} \right)^{\sigma-1} + \left(\frac{1}{w} \right)^{\sigma-1} \right] = \Phi$, this implies $Y = \Phi \Omega$. Combining with the aggregate factor demand functions (55) and (56), we obtain the standard text-book aggregate CES production function,

$$Y = (\Phi) \left(\frac{1}{r} \right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{1}{w} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (58)$$

where (Φ) is the same as before.

6.2 The Cobb-Douglas Function

The above analysis requires $\sigma > 1$ in the Pareto distribution, otherwise the aggregate output in (49) is not well-defined. This is due to the long tail property of the Pareto distribution. In this subsection, we extend the results to the case of $\sigma \in (0, 1]$.

Define a modified production function of (32) as,

$$i = \int_0^1 (i,j) () \quad (59)$$

where (i,j) is a truncated linear function satisfying

$$(i,j) = \begin{cases} i & \text{if } j \leq a_{\max} \\ a_{\max} & \text{if } j > a_{\max} \end{cases} \quad (60)$$

where $a_{\max} \in (0, 1)$ is an arbitrary truncation point. Firm j will demand intermediate good i if and only if $(i,j) > 0$. Given the Pareto distribution $(i) = 1 - \left(\frac{i}{a}\right)^\sigma$, it follows immediately that the aggregate demand for intermediate good i is determined by

$$j = \begin{cases} \Omega \left(\frac{1}{p_j}\right)^\sigma & \text{if } j \leq a_{\max} \\ 0 & \text{otherwise} \end{cases} \quad (61)$$

The aggregate output is determined by $y = \Omega \int_0^1 j$ where $j = \int_{h(a) \geq p_j} (i)^\sigma$ is the expected output per firm by using intermediate good i .

First consider 0 j. 1. Given the definition of (i,j) , we have

$$\begin{aligned} j &= \left[\int_{p_j}^{a_{\max}} (i)^\sigma + \int_{a_{\max}}^\infty a_{\max}^\sigma \right] \\ &= \frac{1-\sigma}{1-a_{\max}^{1-\sigma}} \frac{1}{j} \end{aligned} \quad (62)$$

Note, since $\frac{1-\sigma}{j} = (i)^\sigma \frac{1-\sigma}{a_{\max}^{1-\sigma}}$ and $a_{\max} < 1$, we have $j > 0$ for any i . The aggregate output is then

$$\frac{1-\sigma}{a_{\max}^{1-\sigma}}$$

$$\int_0^1 \frac{1-\sigma}{j} = \frac{1-\sigma}{\max} (1) \Phi \quad (65)$$

Using this relationship and $\Omega = \Omega \left[\frac{a_{\max}^{1-\sigma}}{1-\sigma} \frac{\sigma}{1-\sigma} \int_0^1 \frac{1-\sigma}{j} \right]$ we can establish the relationship,

$$= \left[\Phi + \frac{1-\sigma}{\max} \right] \Omega \quad (66)$$

Recall that $j = \Omega j^{-\sigma}$, so we have $\frac{\sigma-1}{j^{\frac{\sigma-1}{\sigma}}} = \Omega^{\frac{\sigma-1}{\sigma}} j^{1-\sigma}$. Integrating both sides over and utilizing equation (65) gives $\int_0^1 \frac{\sigma-1}{j^{\frac{\sigma-1}{\sigma}}} = \Omega^{\frac{\sigma-1}{\sigma}} \left[\frac{1-\sigma}{\max} (1) \Phi \right]$, which gives a relationship between Ω and $\left[\int \frac{\sigma-1}{j^{\frac{\sigma-1}{\sigma}}} \right]^{\frac{\sigma}{\sigma-1}}$. Substituting this relationship into equation (66), we then obtain an aggregate production function similar to (49):

$$= (\Phi) \left(\int_0^1 \frac{\sigma-1}{j^{\frac{\sigma-1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}} \quad (67)$$

where $(\Phi) = \left(\Phi + \frac{1-\sigma}{\max} \right) \left[\frac{1}{a_{\max}^{1-\sigma} - (1-\sigma)\Phi} \right]^{\frac{\sigma}{\sigma-1}}$ is the TFP. Notice if the total entry cost is less than total output,

$$\Omega \Phi = \frac{\Phi}{\Phi + \frac{1-\sigma}{\max}} \quad (68)$$

then we have $\frac{1-\sigma}{\max} (1) \Phi > 0$, which ensures the TFP and the aggregate output in equation (67) is well-defined.

To derive the Cobb-Douglas function, consider $\sigma = 1$. In this case the distribution becomes $(j) = 1 - \frac{1}{a}$ and we have $j = \Omega \frac{1}{p_j}$ and

$$\begin{aligned} j &= \int_{p_j}^{a_{\max}} -1 + \int_{a_{\max}}^{\infty} \max^{-2} \\ &= \log(\max) - \log(j) + 1 \end{aligned} \quad (69)$$

Hence we have

$$= [\log(\max) + 1] \Omega - \Omega \int_0^1 \log(j) \quad (70)$$

The total cost for input is $\int_0^1 j = \Omega$, and the total profit is $\Pi = \Omega \left[\log(\max) - \int_0^1 \log(j) \right]$.

The free entry condition then becomes $\Phi = \log(\max) - \int_0^1 \log(j)$. Hence, we have

$$\begin{aligned}
&= [\log(\max) + 1] \Omega - \Omega [\log(\max) - \Phi] \\
&= (1 + \Phi) \Omega
\end{aligned} \tag{71}$$

Also, since $\Omega \frac{1}{p_j} = \int_0^1 \log(j) dj = \log(\Omega) - \int_0^1 \log(j) dj$, which by the free entry condition implies $\log(\Omega) - \int_0^1 \log(j) dj = \log(\max) - \Phi$, or alternatively,

$$\Omega = \frac{\max}{\exp(\Phi)} \exp \left\{ \int_0^1 \log(j) dj \right\} \tag{72}$$

The aggregate output is then

$$= \frac{(1 + \Phi) \max}{\exp(\Phi)} \exp \left\{ \int_0^1 \log(j) dj \right\} \tag{73}$$

which is the Cobb-Douglas function with continuum of inputs. Notice the TFP is decreasing in Φ .

A special case of the above example is when the production function of the final good firm is given by $y_i = (k_i)^\alpha + (l_i)^\beta$ where k_i is capital, l_i is labor, and (\cdot) is a truncated linear function defined in (60). The factor demand for capital and labor are given, respectively, by (51) and (52). Similar operations can show that the aggregate production function is given by

$$= \tilde{\omega}(\Phi) \frac{\alpha}{\alpha+\beta} \frac{\beta}{\alpha+\beta} \tag{74}$$

where $\tilde{\omega}(\Phi) = (\max + \Phi) \left(\frac{a_{\max}}{\alpha} \right)^{\frac{\alpha}{\alpha+\beta}} \left(\frac{b_{\max}}{\beta} \right)^{\frac{\beta}{\alpha+\beta}} \exp \left(\frac{-\Phi}{\alpha+\beta} \right)$.

6.3 The Quadratic Function

It is also possible to deduce a quadratic aggregate production function from our model if the Pareto distribution is replaced by the uniform distribution. Consider $y_i = \int_0^1 (k_i)^\alpha (l_i)^\beta$, where k_i is a random draw from the uniform distribution with support $[0, 1]$. The aggregate demand for intermediate good j is

$$j = \begin{cases} (1 - j) \Omega & \text{if } j \in [0, 1] \\ 0 & \text{if } j > 1 \end{cases} \tag{75}$$

The aggregate output is determined by $Y = \frac{\Omega}{2} \int_0^1 (1 - j)^2$. Substituting out j using the aggregate demand relation, $j = 1 - \frac{X_j}{\Omega}$, we have

$$= \int_0^1 \left(1 - \frac{X_j}{\Omega} \right)^2 \tag{76}$$

The profit function is

$$t = \int_0^1 [j(1-j)] = \frac{1}{2} \int_0^1 (1-j)^2 \quad (77)$$

With free entry, we have $t = \Phi$. Combining (77) with the aggregate demand function (75) gives $\Omega = \left(\frac{1}{2\Phi} \int_0^1 j^2 \right)^{\frac{1}{2}}$. Substituting this into (76) gives the quadratic aggregate production function,

$$= \int_0^1 j \left(\frac{\Phi}{2} \right)^{\frac{1}{2}} \left(\int_0^1 j^2 \right)^{\frac{1}{2}} \quad (78)$$

7 Conclusion

In this paper we have proposed a flexible price RBC model with entry and exit to explain the puzzling effects of technology shocks on the economy, especially the asymmetric impacts of aggregate and sector-specific technology shocks on sectorial activity. Key elements of our explanation are net business formation at the aggregate level and factor-demand rigidity at the microlevel. Our model collapses to a standard frictionless one-factor RBC model.

$$\Phi = \frac{t}{t+1} (t_{t+1} + (1 - t_{t+1})\Phi) \quad (\text{A4})$$

$$\Omega_{t+1} = (1 - t_t)\Omega_t + t_t \quad (\text{A5})$$

$$t_t t_t^{\alpha-1} t_t^{1-\alpha} = t_t + \quad (\text{A6})$$

$$(1 - t_t) t_t t_t^{\alpha} t_t^{-\alpha} = t_t \quad (\text{A7})$$

$$t_t t_t^{-1} = (1 - t_t)^{-1} \quad (\text{A8})$$

$$t_t^{-1} = t_t t_{t+1}^{-1} (1 + t_{t+1}) \quad (\text{A9})$$

$$t_t + t_{t+1} (1 - t_t) t_t \Phi = t_t \quad (\text{A10})$$

Notice that only the first three equations are different from the previous model in Section (4).

Define $j = \frac{p(1-F_j)}{G_j}$. Notice that the numerator is the expected cost and the denominator is the expected output of using intermediate good j ; hence, j can be interpreted as the expected marginal cost of intermediate good j . Since $j = \int_p^\infty j(\cdot) \quad \int_p^\infty j(\cdot) = (1 - j)$, we have $j < 1$ and $j > j$ for all j . This implies $\int j > \int j$. The price elasticity of j is given by $\frac{p^2 f_j}{G_j} = \frac{p(1-F_j)}{G_j} \frac{p f_j}{1-F_j} = j j$.

Calibration. For simplicity, we assume that in the steady state, $j(\cdot)$ and $j(\cdot)$ are the same across intermediate goods sector j , but the probability density function $j(\cdot)$ remains heterogenous across j . This is enough to ensure that the sectorial impulse responses to technology shocks remain heterogenous across sectors. Using circumflex to denote a log-linearized variable around its steady state and using $\bar{\cdot}$ to denote $\int j$, log-linearizing the first three equations gives

$$\hat{t}_t + (1 - \bar{t}_t) \hat{t}_t + \hat{t}_t + \bar{t}_t \hat{t}_t = \hat{\Omega}_t \quad (\text{A11})$$

$$\hat{t}_t = \hat{\Omega}_t - \bar{t}_t \hat{t}_t \quad (\text{A12})$$

$$\hat{t}_t = \frac{\int j}{\bar{t}_t} (\bar{t}_t) \hat{t}_t + \frac{\int j}{\bar{t}_t} (1 - \bar{t}_t) \hat{t}_t \quad (\text{A13})$$

Note $\int j = \frac{Y}{\Omega}$. By equation (A4), we have $\Phi = [t^{-1} (1 - t)]\Phi$ in the steady state. The total fixed cost is $\Omega\Phi$ and supposing it accounts for the fraction $\frac{\theta\Omega\Phi}{Y}$ of final output, then we have $(\int j) = \frac{Y}{\Omega\pi} = \frac{\theta\Phi}{\pi\xi} = \frac{\beta\theta}{1-\beta(1-\theta)} \frac{1}{\xi}$. So equation (A13) can be rewritten as

$$\hat{t}_t = \frac{1}{1 - (1 - \bar{t}_t)} \bar{t}_t \hat{t}_t + \left(\frac{1}{1 - (1 - \bar{t}_t)} - 1 \right) (1 - \bar{t}_t) \hat{t}_t \quad (\text{A14})$$

In addition, if we assume without loss of generality that the expected marginal cost equals the equilibrium price of the intermediate good in the steady state, $p_j = p$, then $p_j = 1 - \alpha$ and by equations (A1)-(A3) we have $p = \alpha^{-1} p_j^{1-\alpha}$ and $p = (1 - \alpha) \frac{Y}{\Omega}$. In this case, the steady-state marginal cost is determined by $p = 1 - \frac{\pi\Omega}{Y} = 1 - \frac{\xi(1-\beta(1-\theta))}{\beta\theta}$ and the dynamics of a firm's profit can be further simplified to

$$\hat{\pi}_t = \frac{1}{1} \hat{\pi}_t \quad (\text{A15})$$

Consequently, the heterogenous-firm model has just two additional parameters to calibrate; namely, the steady-state share of entry cost to GDP (ξ) and the average price elasticity across sectors (θ). Let $\xi = 0.01$, $\theta = 0.2$ and the other parameters remain the same as in the previous model; namely, $\alpha = 0.96$, $\beta = 0.1$, $\delta = 0.4$, $\sigma = 0.25$ and the steady state hours worked $\bar{h} = 0.2$. These parameter values imply the steady-state markup is about one percent.¹⁸

¹⁸ Assuming larger markups does not affect our results.

References

- [1] Basu, S., 1998, Technology and business cycles: How well do standard models explain the facts? In *Beyond Shocks: What Causes Business Cycles?* Jeffrey C. Fuhrer and Scott Schuh (ed.), 207–55. Conference Series 42. Boston: Federal Reserve Bank of Boston.

- [14] Francis, N., Ramey, V.A., 2005. Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. *Journal of Monetary Economics* 52(8), 1379-1399.
- [15] Gali, J., 1999. Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review* 89(1), 249-271.
- [16] Gali, J., Lopez-Salido, D., and Valles, J., 2003. Technology shocks and monetary policy: assessing the Fed's performance. *Journal of Monetary Economics* 50(4), 723-743.
- [17] Gali, J., Rabanal, P., 2004. Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar U.S. Data? NBER Working Paper No. 10636.
- [18] Ghironi, F. and M. Melitz, 2005, International trade and macroeconomic dynamics with heterogeneous firms, *The Quarterly Journal of Economics* 120(3), 865-915.
- [19] Houthakker, H.S., 1955-56, The Pareto distribution and the Cobb-Douglas production function in activity analysis, *Review of Economic Studies* 23(1), 27—31.
- [20] Johansen, L., 1972, *Production Functions*, Amsterdam-London: North-Holland.
- [21] King, R., C. Plosser and S. Rebelo, 1988, Production, growth and business cycles: I. The basic neoclassical model, *Journal of Monetary Economics* 21(2-3), 195-232.
- [22] Kydland, F., and E. Prescott, 1982, Time to build and aggregate fluctuations, *Econometrica* 50, 1345-1370.
- [23] Lagos, R., 2006, A model of TFP, *Review of Economic Studies* 73(4), 983-1007.
- [24] Levhari, D., 1968, A Note on Houthakker's Aggregate Production Function in a Multifirm Industry, *Econometrica* 36(1), 151-154.
- [25] Liu, Z. and L. Phaneuf, 2006, Technology shocks and labor market dynamics: Some evidence and theory, *Journal of Monetary Economics* (forthcoming).
- [26] Ramey, V., 2004, Comment on "Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar U.S. Data?" by Jordi Galí and Pau Rabanal, *NBER Macroeconomics Annual 2004*, pp. 307-317.
- [27] Rotemberg, J., Woodford, M., 1996. Real-business-cycle models and the forecastable movements in output, hours, and consumption. *American Economic Review* 86(1), 71–89.

- [28] Vigfusson, R., 2004, The delayed response to a technology shock: A flexible price explanation, Board of Governors of the Federal Reserve System International Finance Discussion Paper 810.