1-1-2014

Firm and Industrial Dynamics Over the Business Cycles

Can Tian
University of Pennsylvania, tiancan@sas.upenn.edu

Firm and Industrial Dynamics Over the Business Cycles

Abstract

This dissertation consists of three essays. In Chapter 1, we proposes a dynamic multi-sector production network model in which firms receive news on the future product-specific demand of a representative household. Since production takes time and firms in the production sectors are connected via input-output links, news on the future final demand of an individual product changes firms' forecasts of their future sales, creating economy-wide effects named as forecast shocks. Forecast shocks are transferred upwards through the supplier-customer connections in the network, from the buyer of an input good to the producer. The model explains the asymmetry in the transmission of individual shocks in the network and how shocks to the expectations generate real, persistent effects. The equilibrium is analytically solved and calibrated to the U.S. economy. Quantitative analysis then follows to examine the model performance. In Chapter 2, we incorporate a firm's project choice decision into a firm dynamics model with business cycle features to explain this empirical finding both qualitatively and quantitatively. In particular, all projects available have the same expected flow return and differ from one another only in the riskiness level. The endogenous option of exiting the market and limited funding for new investment jointly play an important role in motivating firms' risk-taking behavior. The model predicts that relatively small firms are more likely to take risk and that the cross-sectional productivity dispersion, measured as the variance/standard deviation of firm-level profitability, is larger in recessions. In Chapter 3, we consider the impact of job rotation in a directed search model in which firm sizes are endogenously determined, and match quality is initially unknown. In a large firm, job rotation allows the firm to at least partially ameliorate losses from mismatches of workers to jobs. As a result, in the unique equilibrium, large firms have higher labor productivity and lower separation rate. In contrast to the standard directed search model with multi-vacancy firms, this model can generate a positive correlation between firm size and wage without introducing exogenous productivity shocks or a non-concave production function.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Dirk Krueger

Subject Categories
Economics

FIRM AND INDUSTRIAL DYNAMICS OVER THE BUSINESS CYCLES

Can Tian

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

____

Dirk Krueger, Professor of Economics

Graduate Group Chairperson

____

George J. Mailath, Walter H. Annenberge Professor in the Social Science

Dissertation Committee

Harold L. Cole, Professor of Economics

Iourii Manovskii, Associate Professor of Economics

FIRM AND INDUSTRIAL DYNAMICS OVER THE BUSINESS CYCLES

To Fei Li

ABSTRACT

FIRM AND INDUSTRIAL DYNAMICS OVER THE BUSINESS CYCLES

Can Tian

Dirk Krueger

This dissertation consists of three essays. In Chapter 1, we proposes a dynamic multi-sector production network model in which firms receive news on the future product-specific demand of a representative household. Since production takes time and firms in the production sectors are connected via input-output links, news on the future final demand of an individual product changes firms' forecasts of their future sales, creating economy-wide effects named as forecast shocks. Forecast shocks are transferred upwards through the supplier-customer connections in the network, from the buyer of an input good to the producer. The model explains the asymmetry in the transmission of individual shocks in the network and how shocks to the expectations generate real, persistent effects. The equilibrium is analytically solved and calibrated to the U.S. economy. Quantitative analysis then follows to examine the model performance. In Chapter 2, we incorporate a firm's project choice decision into a firm dynamics model with business cycle features to explain this empirical finding both qualitatively and quantitatively. In particular, all projects available have the same expected flow return and differ from one another only in the riskiness level. The endogenous option of exiting the market and limited funding for new investment jointly play an important role in motivating firms' risk-taking behavior. The model predicts that relatively small firms are more likely to take risk and that the cross-sectional productivity dispersion, measured as the

variance/standard deviation of firm-level profitability, is larger in recessions. In Chapter 3, we consider the impact of job rotation in a directed search model in which firm sizes are endogenously determined, and match quality is initially unknown. In a large firm, job rotation allows the firm to at least partially ameliorate losses from mismatches of workers to jobs. As a result, in the unique equilibrium, large firms have higher labor productivity and lower separation rate. In contrast to the standard directed search model with multi-vacancy firms, this model can generate a positive correlation between firm size and wage without introducing exogenous productivity shocks or a non-concave production function.

Contents

List of Tables

List of Figures

Chapter 1

Forecast Shocks in Production Networks

1.1 Introduction

A key theme of macroeconomics is the search for the causes of aggregate economic fluctuations. While many model the driving force as shocks to aggregate fundamentals, a large branch of literature, pioneered by Long and Plosser (1983) and Jovanovic (1987), attributes the observed aggregate fluctuations to a micro origin. Idiosyncratic shocks at the firm- or industry-level can explain the observed aggregate economic fluctuation to a large extent. However, this literature typically overlooks the non-trivial role of the topology of the customer-supplier network in determining the direction of the transmission of idiosyncratic shocks. For example, Shea (2002) and Conley and Dupor (2011) show that, at the sectorial level, the demand-side linkage is more important. More recently, Kelly, Lustig, and Van Nieuwerburgh (2013) suggest that firm volatility data favor firm-level shock transmission in the customer-to-supplier direction over the opposite.

This paper introduces forecast shocks into the dynamic framework of Long and Plosser (1983) and addresses the paths of shock transmission. In particular, I assume that in each period, agents in the economy observe a common external signal (news) on the consumer's preference for different products in the future. Firms forecast their sales relative to the total value added in the following period. Since each sector has to decide its inputs one period ahead, in equilibrium, the formation of their forecasts has a recursive structure and every current forecast in fact summarizes future forecasts. Once agents in the economy at period t_0 observe the news and anticipate a change in forecast by sector i in the future at period $t_0 + T$, such anticipation will immediately show up in the formation of current-period forecasts by relevant sectors. Specifically, anticipating the change in forecast by sector i at t_0+T, i's direct upstream sectors (suppliers) will change their forecasts accordingly at t_0+T-1, which in turn leads the suppliers' suppliers to adjust their forecasts at $t_0 + T - 2$, and so forth. Hence, a path of transmission of fluctuation from the downstream sectors to the upstream sectors is constructed. Notice that in my model, once external news of the demand arrives, all agents in the economy receive such information. As a result, there is no heterogenous information issue in my model.

The input-output network plays a dual role in this model. In the formation of forecasts, shocks to future forecasts are transferred upwards through the supplier-customer connections in the network, from the buyer of an input good to the producer. The shares of industrial sales then reflect the updated forecasts, which have real effects on levels of output, consumption, input, etc. There is also a conventional

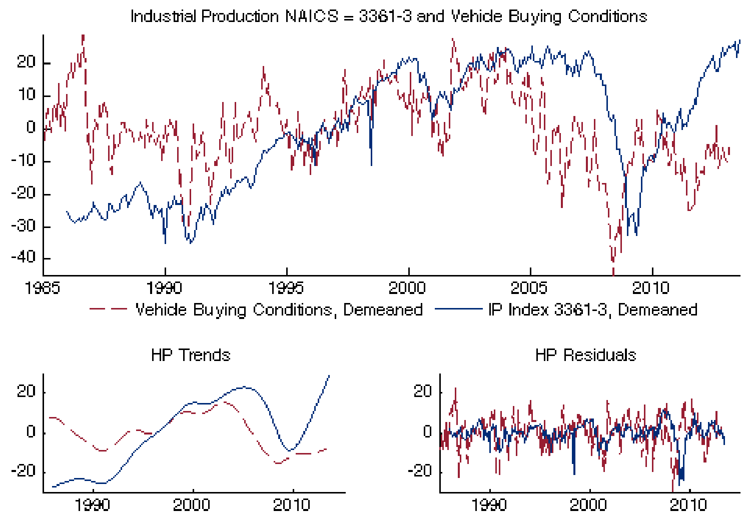

Figure 1.1: Vehicle Buying Conditions (VBC) and Industrial Production (IP) of NAICS sectors 3361-3. The top panel shows the raw, demeaned data of the two sequences and the bottom two panels are the HP trends and residuals with monthly smoother 14400. Red line: monthly VBC from Jan 1985 to Mar 2013. Blue line: monthly IP 3361-3 from Jan 1986 to Aug 2013.

role in the transmission of real effects. Any shock that changes the current-period output of a sector will have a prolonged effect on the economy through the one-period-ahead input choices downwards to its customer sectors. This is also how productivity shocks are transferred.

This paper treats the news on demand as a source of aggregate volatility. In particular, I focus on the shocks to the expectation of product-specific demand which can lead to changes in real economic activities. In general, the future demand of households depends on many factors such as their income, the price of the product, their expectations of the prices of the product in the future, and fiscal policy. Consider motor vehicle production as an example. As part of the Survey of Consumers conducted by University of Michigan, the time series of indices of Buying Condi-

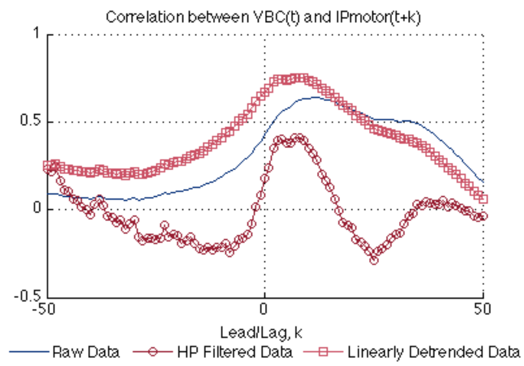

Figure 1.2: Cross-Correlogram: VBC(t) versus IP 3361-3 $(t-50, t+50)$. The correlation coefficients achieve their maximum around k=10. This indicates that the VBC leads the production.

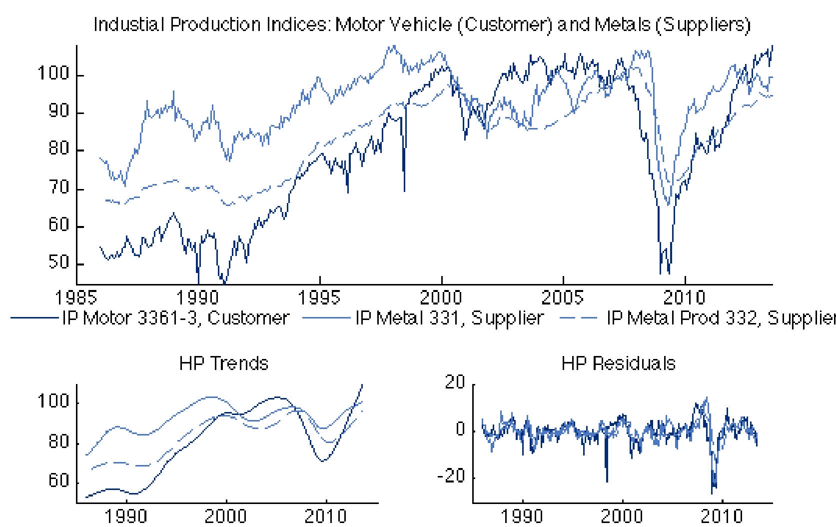

Figure 1.3: Industrial Production: NAICS Sectors 3361-3, 331, and 332 from Jan 1986 to Aug 2013. The top panel shows the raw data of the three sequences and the bottom two panels are the HP trends and residuals with monthly smoother 14400.

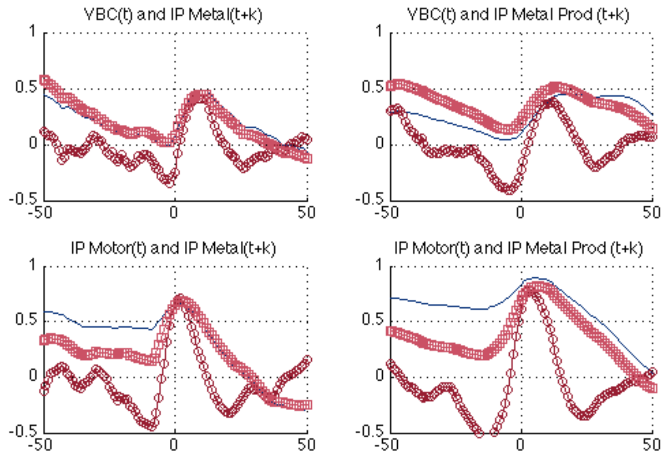

Figure 1.4: Cross-Correlogram: VBC(t) versus IP 331, IP 332 $(t-50, t+50)$ and IP 3361-3 (t) versus IP 331, IP 332 $(t-50, t+50)$.

tions for Vehicles (henceforth VBC) show the consumers' willingness to by and/or the economic conditions of buying household motor vehicles in the following year.[1] I interpret this indicator as the expected demand for motor vehicles. Figure 1.1 plots the co-movement between the monthly VBC and the monthly index for industrial production of NAICS sectors 3361-3, producers of motor vehicles. The time series data of indices for industrial production are part of the G.17 series, calculated and released by the Board of Governors of the Federal Reserve System. The VBC, or the expected demand, seems to lead the output of motor vehicles when the data are in their raw form, HP filtered, or even in the trend. Reporting the cross-correlogram between these two series, Figure 1.2 confirms the lead/lag relationship. I then turn to the major suppliers of inputs required by vehicle producers. In fact, around

[1]In the survey, consumers are asked the following question: "Speaking now of the automobile market – do you think the next 12 months or so will be a good time or a bad time to buy a vehicle, such as a car, pickup, van, or sport utility vehicle?" Among the consumers who feel it will be a good time to buy a vehicle, low loan interest rates and low price expectation are among the main reasons cited for buying one.

38.8% of the input expenditure of this sector is spent on its own products, 9.3% on fabricated metal products, and 8.6% on primary metals, which are products of NAICS sectors 332 and 331, respectively. Not surprisingly, the suppliers' output levels comove with that of the customer, as shown by Figure 1.3. Additionally, the top two panels of Figure 1.4 depict the lead/lag relationship between VBC and the output of each supplier sector. Figure 1.4 shows a similar pattern to Figure 1.2. This evidence supports my claim that the expectation regarding the demand for a certain product, in this example, vehicles, has real impact not only on the producer but also on the producer's input suppliers. The bottom two panels of Figure 1.4 show that the outputs of the two supplier sectors also comove with that of the customer sector. Meanwhile, the positive relationship between the lagged output of the suppliers and the current output of the customer, albeit weak, suggests a timing difference between the input purchase and the production, as assumed by Long and Plosser (1983).

The forecast shocks feature upward transmission direction, contrary to the downward transmission of productivity shocks in the same model. In reality, one expects the shock transmission through the supplier-customer links to move in both directions, instead of uni-directionally. However, recent empirical studies provide evidence that suggests the upward direction from the customer to the supplier is more important. Estimating structural equilibrium models, Shea (2002) finds that demand-side linkage is important in generating output comovement at the sector level, hence the upward transmission of sector-specific shock is important at the aggregate level. Conley and Dupor (2003) find the strongest evidence for complementarity when sectors are "close" to each other according to a distance measure that captures Shea's demand-side linkage. At the firm level, Kelly, Lustig and Van Nieuwerburgh (2013) suggest that upstream shock propagation provides a better description of firm volatility data than downstream. Among publicly traded firms, Cohen and Frazzini (2008) find evidence of "customer momentum", that is, predictable stock return for the supplier firm when there are shocks to its linked customer firms. Additionally, they show that present customer shocks have significant predictability over the supplier's future real activity while the predictability does not exist without the link.

In the main model, I assume that the source of the forecast fluctuation is driven by news shocks based on the consumers' product-specific preference in the future. The news is received by all agents in the economy, so there is no asymmetric information across agents. I solve the model analytically and use the U.S. Bureau of Economic Analysis (BEA) Annual Industrial Accounts to calibrate the model and quantitatively illustrate the importance of forecast dynamics. Using the calibrated model, I compute the sector-specific forecast sequences from the standard use tables between 1997 and 2012. Under the assumptions that final consumption shares follow the Dirichlet distribution and that the news follows the multinomial distribution, I estimate the parameters for the processes and simulate the model. The quantitative analysis shows that, without any productivity shock, the model can generate non-trivial fluctuations in the economic activity both at the aggregate level and at the industry level. The model shows limited success in capturing the comovement between industries. The news also explains the positive relationship between the input prices and input uses. Last but not least, the model demonstrates how volatility in the aggregate productivity measured as Solow residual can be observed even though no productivity shock is present.

1.1.1 Related Literature

First, my paper is directly related to the multi-industry real business cycle models. There is a longstanding debate: Can idiosyncratic *productivity* fluctuations cause business cycles at the aggregate level. To the best of my knowledge, Long and Plosser (1983) develop the first model to study this issue. Horvath (1998, 2000) and Dupor (1999) introduce capital accumulation into Long and Plosser (1983). However, this literature ignores that the aggregation of idiosyncratic shock critically depends on the topology of the input-output network. Hence, there is an emerging literature that studies the role of the topology of the input-output network in macroeconomics. Acemoglu et al. (2012) consider a static multi-sector model and study the role of the input-output network in the aggregation of idiosyncratic productivity shock in different sectors. They show that independent idiosyncratic shocks in different sectors cannot offset each other when the network is asymmetric. In contrast to the previous literature, my paper studies news shock instead of productivity shock. In Long and Plosser (1983)'s framework, the productivity shock does not affect the share of sales. However, the share changes over time, which indicates the existence of another source of fluctuation. I introduce sector-specific forecast shock. I also study how the network structure determines the transmission mechanism of sector-specific forecast shocks. Unlike Horvath (1998, 2000) and Dupor (1999), the dynamics in my model come purely from the information and time-to-produce mechanism instead of from capital accumulation. In addition, in my model, the forecast shock is transmitted from the downstream sectors to the upstream sectors, which is consistent with the recent firm-level empirical studies by Kelly, Lustig and Van Nieuwerburgh (2013).

The comovement between industries is a feature of the business cycles and is in itself a crucial topic with in the multi-industry literature, e.g., Shea (2002), Conley and Dupor (2003), Foerster et al. (2011), and recently Atalay (2013). While most of the papers focus on the importance of the sectoral links in propagating productivity shocks, I focus on the shocks originating from the demand side. In fact, the quantitative analysis in this paper shows that positive relationship in economic activity between industries can come from cross-sectionally negatively correlated news. Additionally, while Atalay (2013) interprets the fluctuation in an industry's input expenditure share and its positive correlation with input prices as an evidence for non Cobb-Douglas production technology, I demonstrate in the quantitative analysis that the forecast shocks can be another explanation.

Second, in my paper, shocks come from the demand side, either the consumer's contemporaneous preference or the news about future preference. Hence, my paper is also related to the notion of demand side driven business cycles. Baxter and King (1991) first introduced demand shocks into a neoclassical framework with a representative production sector, in which the demand shocks can partially explain the U.S. business cycles in the presence of increasing returns to scale technology and/or productivity shocks.

Third, my paper clearly relates to the news literature, see Lorenzoni (2011) for a detailed survey. In this literature, one assumes that the consumers and firms receive expectation shock on the technology in the future, and studies how the news shock affects the demand and output. In my model, however, the functional assumption on the production functions and the utility function prevents the productivity news from having any effect, in order to highlight the roles of news about future preference.

The rest of this paper is organized as follows. In section 1.2, I present the model and its equilibria. In section 1.3, I discuss the dynamics of forecast in the equilibrium. In section 1.4, I calibrate the model and quantitatively explore the importance of forecast shock. Section 1.5 concludes. The appendix contains the omitted proofs and detailed description of the data and the estimation process.

1.2 Model

1.2.1 Setup

I consider a neoclassical multi-sector model following Long and Plosser (1983). Time is discrete with infinite horizon, $t = 0, 1, 2, ...$ The economy consists of n competitive industries denoted by $\{1, 2, ..., n\}$, each of which produces a distinct type of good. Each good can be consumed by consumers or used as an input for the production of other goods in the following period.

Firms. There is a representative firm in each of the n industries. At time t, the production of good i by industry i requires labor hired at t and a variety of goods as inputs, the amount of which is determined in the previous period. Each firm employs a time-invariant Cobb-Douglas production technology with constant returns to scale. In addition, the production at each sector is subject to some idiosyncratic productivity shock. Specifically, the technology of industry i transforms h_{it} units of labor and x_{ijt-1} units of pre-determined amount of good j, $\forall j = 1, ..., n$, into y_{it}

units of output, determined by

$$y_{it} = z_{it} h_{it}^{1-\alpha_i} \prod_j x_{ijt-1}^{\alpha_i \omega_{ij}}$$

where z_{it} is the realized productivity term. Define the productivity vector at time t as $\mathbf{z}_t = (z_{1t}, ..., z_{nt})'$, which is drawn from a stationary process,

$$\mathbf{z}_t \sim \Xi^z (\cdot | \mathbf{z}_{t-1})$$

with uncoditional mean $\mathbf{z} = (z_1, ..., z_n)$

$$E(\mathbf{z}_t) = \mathbf{z}.$$

$\alpha_i \in (0, 1)$ is i's total share of input use and $1 - \alpha_i$ is the labor share. Note that, in absence of capital, labor is the only value-added input for each industry. Out of i's total input use, the share of j's product as input is $\omega_{ij} \geq 0$, which captures the importance of good j in producing i. When $\omega_{ij} > 0$, industry i is a customer of good j and industry j is a supplier. The constant returns to sale technologies require that $\sum_j \omega_{ij} = 1$ for each industry i. Define

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix}, \Omega = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \vdots & \vdots & & \vdots \\ \omega_{n1} & \omega_{n2} & \cdots & \omega_{nn} \end{bmatrix}$$

Ω determines the production architecture, which is in fact a directed network amongst all industries with weighted links, in which the direction is simply that of the flow of input goods. The importance of a link from industry j to i is captured by the value of ω_{ij}. Generically, Ω is an asymmetric matrix.

In each period t, firm i receives flow profit π_{it}, which is the sales of good i minus the labor compensation and the input purchase for production in the following period:

$$\pi_{it} = p_{it}y_{it} - w_{it}h_{it} - \sum_j p_{jt}x_{ijt}.$$

The GDP produced by this industry is equivalent to the value-added generated by this industry, which consists of the profit π_{it} and the labor compensation $w_{it}h_{it}$.

Consumers. In addition to the firms, there is a representative long-lived household that gains utility from consuming a variety of goods and supplies labor to each sector. The preference (flow payoff) is modeled as

$$\eta_t \sum_i \theta_{it} u\left(c_{it}\right) - \sum_i v\left(h_{it}\right),$$

where η_t is the aggregate preference parameter, θ_{it} governs the good-i-specific preference, c_{it} is the amount of good i consumed at time t, and h_{it} is the labor supplied to the firm in industry i at t. Assume the preference parameter $\eta_t > 0$ follows a Markovian process such that

$$\eta_{t+1} \sim \Xi^\eta\left(\cdot|\eta_t\right), \text{ and } E\left(\frac{\eta_{t+1}}{\eta_t}\right) = E\left(\frac{\eta_{t+1}}{\eta_t}|\eta_t\right) = 1.$$

The utility takes the logarithm form, $u\left(c\right) = \ln c$, and the disutility of working, $v\left(h\right) = h^{1+\varepsilon}/\left(1+\varepsilon\right)$, $\varepsilon \geq 0$. Let vector θ_t summarize product-specific preference parameters at time t such that $\theta_t = \left(\theta_{1t}, \theta_{2t}, ..., \theta_{nt}\right)'$ and $\sum_{i=1}^n \theta_{it} = 1$, $\theta_{it} \geq 0$.

Assume that nature draws all $\{\theta_t\}_{t=0}^{+\infty}$ at the beginning of time while all agents in the economy share a common prior belief that θ_t is independent and identically distributed over time, with the expectation $\theta = (\theta_1, ..., \theta_n)'$, and $\sum_{i=1}^{n} \theta_i = 1$, $\theta_i \geq 0$, $\forall i$,

$$\theta_t \sim \Xi^{\theta}(\cdot), \text{ and } E(\theta_t) = \theta$$

As will be shown later, θ_{it} is the equilibrium share of consumption expenditure on good i at time t. The household receives labor income and all profits and discounts the future at rate $\beta \in (0, 1)$. To ensure the uniqueness of the equilibrium that is defined in due course, the following assumption on parameter matrices invertibility is required.

Assumption 1. $(I - \beta\Omega'\mathbf{A})$ *is invertible, where* I *is the identity matrix. A sufficient condition is that* $\lim_{k\to\infty} (\beta\Omega'\mathbf{A})^k = 0$.

Information Structure. Assume that at any time t, all agents share a common information set $\mathcal{I}_t$. The economy evolves according to the following timeline. At the beginning of period t, firms in each sector inherit from the previous period the input goods, $\{x_{ijt-1}\}_{j=1}^{n}$ for all sector i. All shocks to fundamentals of the current period are realized and become observable, including $(\{z_{it}, \theta_{it}\}_{i=1}^{n}, \eta_t)$. At the same time, agents in this economy receive a set of signals M_t from an exogenous source and the information set updates such that $\mathcal{I}_t = \mathcal{I}_{t-1} \cup \{\{z_{it}, \theta_{it}\}_{i=1}^{n}, \eta_t; M_t\}$. I assume the signals contain information about future product-specific demand. The specific form of the signals M_t will be discussed with greater detail in due course. However, it is worth noting that M_t is commonly known by all parties in the economy, so there is no heterogenous information in my model. Seeing the wages and prices, firms of all sectors make employment choices and the household provides labor to each sector. Production takes place. Firms decide on input purchase for production in the next period while the household buys the basket of goods for consumption.

1.2.2 Decisions and Equilibrium

Household's Choice. At each time t, after the realization of η_t and θ_t, the household takes wages $\{w_{it}\}_i$, prices $\{p_{it}\}_i$, and profits $\{\pi_{it}\}_i$ as given and chooses labor supply $\{h_{it}\}$ and consumption bundle $\{c_{it}\}$ subject to the budget constraint:

$$\sum p_{it}c_{it} \leq \sum w_{it}h_{it} + \sum \pi_{it}. \tag{1.1}$$

The equality of the constraint holds in equilibrium, equating the total consumption expenditure on the left hand side to the total value added on the right hand side. Therefore, under the model specification, the aggregate GDP is the same as the aggregate consumption expenditure. Let λ_t be the Lagrangian multiplier of this constraint, then household maximization yields the following first order conditions:

$$\beta^t h_{it}^{\varepsilon} = \lambda_t w_{it} \tag{1.2}$$

$$\beta^t \eta_t \theta_{it} = \lambda_t p_{it} c_{it.} \tag{1.3}$$

It is convenient to define the consumption index C_t and the ideal price index P_t^C such that $P_t^C C_t = \sum_i p_{it}c_{it}$, [2]

$$C_t = \prod_i c_{it}^{\theta_{it}}, P_t^C = \prod_i \left(\frac{p_{it}}{\theta_{it}}\right)^{\theta_{it}}.$$

[2]See Appendix A1 for the derivation.

Given price $\{p_{it}\}$ and wage $\{w_{it}\}$, the household's demand $\{c_{it}\}$ and labor supply $\{h_{it}\}$ can be rewritten as follows.

$$p_{it}c_{it} = \theta_{it}P_t^C C_t \tag{1.4}$$

$$h_{it}^{1+\varepsilon} = \eta_t \frac{w_{it}h_{it}}{P_t^C C_t} \tag{1.5}$$

$$\lambda_t = \beta^t \frac{\eta_t}{P_t^C C_t} \tag{1.6}$$

Firms' Choice. At each time t, after the realization of z_{it}, firm i chooses labor demand l_{it} for current production and buys inputs $\{x_{ijt}\}_j$ for future production to maximize profit, facing wage w_{it} and prices $\{p_{jt}\}_j$ and expecting a discount factor $\Lambda_{t,t+1}$. Define firm i's sales at time t as

$$s_{it} = p_{it}y_{it}. \tag{1.7}$$

Firm i's labor demand is given as follows:

$$w_{it}h_{it} = (1-\alpha_i)\, s_{it}.$$

Firm i also needs to decide on the expenditure on each input good j, x_{ijt}, based on the expectation on its next-period sales,

$$p_{jt}x_{ijt} = \alpha_i\omega_{ij}E_t\left(\Lambda_{t,t+1}s_{i,t+1}\right). \tag{1.8}$$

Define the total input expenditure of firm i at time t as u_{it},

$$u_{it} = \sum_j p_{jt} x_{ijt}. \tag{1.9}$$

Firms form expectations based on the common information set at time t, $\mathcal{I}_t$, which summarizes all previous signals $\{M_\tau\}_{\tau \le t}$. Denote $E_t(\cdot) = E(\cdot|\mathcal{I}_t)$. Given the price $\{p_{jt}\}$ and the current information set $\mathcal{I}_t$, one can rewrite firm i's input demand $\{x_{ijt}\}$ as follows:

$$p_{jt} x_{ijt} = \omega_{ij} u_{it} \tag{1.10}$$

$$u_{it} = \alpha_i E_t \left(\Lambda_{t,t+1} s_{i,t+1} \right). \tag{1.11}$$

At equilibrium, a firm's discount factor is consistent with the household's intertemporal concern,

$$\Lambda_{t,t+1} = \frac{\lambda_{t+1}}{\lambda_t} = \beta \frac{\eta_{t+1}}{\eta_t} \frac{P_t^C C_t}{P_{t+1}^C C_{t+1}}. \tag{1.12}$$

Hence, combining equations (1.11) and (1.12) yields the following Euler equation:

$$\frac{u_{it}}{P_t^C C_t} = \beta \alpha_i E_t \left(\frac{\eta_{t+1}}{\eta_t} \frac{s_{i,t+1}}{P_{t+1}^C C_{t+1}} \right). \tag{1.13}$$

The Euler equation establishes each firm's inter temporal decision. The left hand side is the total input expenditure made by industry i at time t relative to the aggregate value added at that time, where the right hand side is the expected and discounted gain from next period's sales by making that expenditure. It is convenient to define firm i's one-period-ahead forecast at t as

$$\tilde{f}_{it} = E_t \left(\frac{\eta_{t+1}}{\eta_t} \frac{s_{i,t+1}}{P_{t+1}^C C_{t+1}} \right). \tag{1.14}$$

Notice that firms form their expectations based on a common information set that includes previous and current news $\{M_\tau\}_{\tau \leq t}$, so the arrival of news in each period affects firms' decisions by changing their expectations.

Formally, the equilibrium concept is defined as follows.

Definition 1. *A Competitive Equilibrium consists of a list of wages and prices* $\{\{w_{it}, p_{it}\}_{i=1}^n\}_{t=0}^\infty$, *allocations* $\left\{\left\{h_{it}, c_{it}, (x_{ijt})_{j=1}^n, y_{it}\right\}_{i=1}^n\right\}_{t=0}^\infty$, *associated with forecasts* $\left\{\left\{\tilde{f}_{it}\right\}_{i=1}^n\right\}_{t=0}^\infty$ *and information sets* $\{\mathcal{I}_t\}_{t=0}^\infty$, *such that for each period* t, *(1) agents form their forecasts* $\left\{\tilde{f}_{it}\right\}_{i=1}^n$ *based on* $\mathcal{I}_t$, *(2) household optimize given prices and firms optimize given prices and* $(x_{ijt-1})_{i,j=1}^n$, *(3) wages clear all labor markets, (4) product prices clear goods markets,* $c_{it} + \sum_j x_{jit} = y_{it}$, $\forall i$, *and (5) information set evolves based on the realizations of exogenous processes and the law of motions of forecasts are consistent with agents' optimal choices and markets-clearing conditions.*

1.2.3 Equilibrium Analysis

The model equilibrium has analytical solutions. I show in this subsection that on the equilibrium path, the forecasts shape the decisions of the agents. Importantly, for each firm representing each industry, the forecasts of its customer industries crucially determine its action.

Proposition 1. *On the equilibrium path, at any time* t, *given* $\left\{\tilde{f}_{jt}\right\}_j$,

1. *the ratio of industry i's total input expenditure to the aggregate value added is determined by*

$$\frac{u_{it}}{P_t^C C_t} = \beta \alpha_i \tilde{f}_{it}, \tag{1.15}$$

2. *the ratio of industry i's total sales to the aggregate value added is determined by*

$$\frac{s_{it}}{P_t^C C_t} = \theta_{it} + \beta \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt}, \forall i. \tag{1.16}$$

The proof can be found in the appendix. Equation (1.15) is simply the Euler equation (1.13) for each industry i, which states that the input expenditure depends on the forecast of sales in the next period. According to equation (1.16), the revenue of sector i from the sales of its product i depends on two parts: final consumption of product i by the household and i's customer sectors' forecasts of their future sales. Since production takes time, the forecast of a customer j's sales in the next period determines j's use of product i as input purchased in the current period. Therefore, for the supplier of a product, the customers' forecasts matter. For the purpose of GDP accounting, it is also useful to define the value added v_{it} generated by each industry i at time t,

$$v_{it} = s_{it} - u_{it} = \pi_{it} + w_{it} h_{it}. \tag{1.17}$$

Then it is straightforward to see that the total value added V_t is the same as the total consumption expenditure by the household, $V_t = \sum_i v_{it} = P_t^C C_t$, and that each industry i contributes the following fraction to the aggregate GDP,

$$\frac{v_{it}}{P_t^C C_t} = \theta_{it} + \beta \left(\sum_j \alpha_j \omega_{ji} \tilde{f}_{jt} - \alpha_i \tilde{f}_{it} \right), \forall i. \tag{1.18}$$

Given the forecasts, the real equilibrium allocations are uniquely pinned down. The labor hired by sector i is determined by

$$h_{it} = \left[(1 - \alpha_i) \eta_t \left(\theta_{it} + \beta \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt} \right) \right]^{\frac{1}{1+\varepsilon}},$$

where the forecasts of the customer sectors future sales have a real effect in the supplier sector's employment decision. Consequently, the current level of labor then determines real output of sector i given the inputs purchased in the previous period

$$y_{it} = z_{it} h_{it}^{1-\alpha_i} X_{it-1}^{\alpha_i},$$

where the input index X_{it-1} is defined as $X_{it-1} = \prod_j x_{ijt-1}^{\omega_{ij}}$. The final consumption of product i by the household is a fraction of sector i's output,

$$c_{it} = y_{it} \frac{\theta_{it}}{\theta_{it} + \beta \sum \alpha_j \omega_{ji} \tilde{f}_{jt}},$$

and i's customer sector j gets x_{jit} as input,

$$x_{jit} = y_{it} \frac{\beta \alpha_j \omega_{ji} \tilde{f}_{jt}}{\theta_{it} + \beta \sum \alpha_k \omega_{ki} \tilde{f}_{kt}}.$$

The forecasts also determine the cross-sectional distribution of industrial sales:

$$\frac{s_{it}}{\sum_j s_{jt}} = \frac{\theta_{it} + \beta \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt}}{1 + \beta \sum_j \alpha_j \tilde{f}_{jt}}. \tag{1.19}$$

The price of good i relative to good j is also set, as $p_{it}/p_{jt} = (s_{it}/y_{it}) / (s_{jt}/y_{jt})$, and so is the case for relative wages, $w_(it)/w_{jt}$. To normalize the prices level, the ideal consumption price index is set to be one in each period, $P_t^C = 1$. Hence, the real aggregate GDP in each period is actually the total consumption index, C_t.

Now we turn to the forecasts. Define the combined vector of forecasts at time t as $\tilde{\mathbf{f}}_t$ such that $\tilde{\mathbf{f}}_t = \left(\tilde{f}_{1t}, ..., \tilde{f}_{nt}\right)$. The following theorem establishes the recursive formation of the one-period-ahead forecasts.

Theorem 1. *The equilibrium forecast of each sector summarizes the expectation of its future share in consumption and the future forecasts of its customer sectors:*

$$\tilde{\mathbf{f}}_t = E_t\left(\theta_{t+1}\right) + \beta \Omega' \mathbf{A} E_t\left(\tilde{\mathbf{f}}_{t+1}\right). \tag{1.20}$$

Proof. Combining (1.14) and (1.16) yields:

$$\begin{aligned} \tilde{f}_{it} &= E_t\left(\frac{\eta_{t+1}}{\eta_t}\left(\theta_{it+1} + \beta \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt+1}\right)\right) \\ &= E_t \theta_{it+1} + \beta \sum_j \alpha_j \omega_{ji} E_t\left(\tilde{f}_{jt+1}\right). \end{aligned} \tag{1.21}$$

The second equality holds because of the assumption that, conditional on $\mathcal{I}_t$, the change in the aggregate preference parameter η_{t+1}/η_t is independent of both $\{\theta_{i\tau}\}_{\tau>t}$ and future $\{\eta_{\tau+1}/\eta_\tau\}_{\tau>t}$, and therefore, η_{t+1}/η_t is also independent of future forecasts $\left\{\tilde{f}_{i\tau}\right\}_{\tau>t}$. □

The theorem states that, for any sector i, at any time t, the equilibrium one-period-ahead forecast $\tilde{f}_{it}$ has a recursive feature in that it summarizes sector i's beliefs about future forecasts by other sectors. In fact, the impact of j's future forecast on i is weighted by the importance of j's input use of good i, that is, the production function parameter $\alpha_j \omega_{ji}$. In other words, the more industry j's production relies on good i as an input, the more industry i values its belief about j's forecasts.

Note that the productivity shocks play no role in the formation of forecasts. Under the functional-form assumptions of the household's preference and the production technologies, both of which are in the Cobb-Douglas forms, the effects of productivity shocks on allocations and on prices cancel each other out completely. Therefore, the forecast of a sector's future sales summarize only the expectation of future demands of its product, by the household and by the customer sectors. This also eliminates the potential room for signals about future productivities.

Comparative Statics. Here I analyze some simple comparative statics with respect to sector-specific productivity and aggregate demand shock. First of all, the productivity term of sector i at time t, z_{it}, has a direct impact on its output, y_{it}, on the consumption of good i, c_{it}, on the price of good i, p_{it}, and on other industries' intermediary use of good i, x_{jit}, $\forall j$. Other conditions held equal, an increase in z_{it} leads to increases in y_{it}, c_{it}, and x_{jit}, $\forall j$ such that $\omega_{ji} > 0$. It also leads to a decrease in p_{it}. Because of the assumptions on the functional forms of the household preference and the production technology of the firms, the impacts of any change in productivity in prices $\{p_{it}\}$ and in levels $\left\{ y_{it}, c_{it}, \{x_{jit}\}_j \right\}$ completely cancel each other out. Hence, changes in productivity do not show in the price-adjusted terms, that is, sales $\{s_{it}\}$, consumption expenditures $\{p_{it}c_{it}\}$, and input purchases $\{p_{it}x_{jit}\}$. Moreover, a change in the productivity of a specific sector propagates to other sectors through the input choices in a downstream direction, traveling only from the producer of a certain product to its immediate customers and then to these customers' customers. Second, while similar in direction to the productivity term, a change in η_t affects all industries as well as the household. Furthermore, the scale of the impact differs from industry to industry or good to good.

1.3 Discussion of Dynamics

1.3.1 Stationary Forecasts

As a benchmark, it is worthwhile to characterize the equilibrium path on which there is no external signal, $M_t = \varnothing$, $\forall t$. Under the current assumptions on the stochastic processes, the lack of further information ensures a time-invariant sequence of forecasts, $\tilde{\mathbf{f}}_t = \tilde{\mathbf{f}}$, where $\tilde{\mathbf{f}} = \left(\tilde{f}_1, ..., \tilde{f}_n\right)'$. The form of the stationary forecasts is established in the following corollary to Theorem 1.

Corollary 1. *The stationary forecasts satisfy the following equation,*

$$\tilde{\mathbf{f}} = \left(I - \beta\Omega'\mathbf{A}\right)^{-1}\theta$$

then $\tilde{\mathbf{f}}$ *is the unique set of time-invariant forecasts on the stationary equilibrium* *path on which $M_t = \varnothing$, $\forall t$. Equivalently written,*

$$\tilde{f}_i = \theta_i + \beta \sum_j \alpha_j \omega_{ji} \tilde{f}_j.$$

Note that, because of the assumptions on Cobb-Douglas production technologies and utility function, any productivity shock that can change z_{it} (which in turn changes y_{it}) is fully absorbed in prices. The case in which $\{z_{it}\}$ is the only source of variation is the one studied by Long and Plosser (1983). In the absence of forecast shocks, the distribution of sales across sectors $\{s_{it}\}$ is constant over time regardless of sector-specific productivity shocks or shocks to the news on future productivity.

A more special case is the steady state of the economy when all processes are set at the determinant mean levels, namely, $\mathbf{z}_t = mathbf z$, $\eta_t = \eta$, and $\theta_t = \theta$, $\forall t$. The steady state will serve as the starting point in the quantitative exploration.

Proposition 2. *The equilibrium outcome at the steady state of the economy consists of $\{h_i,\, y_i,\, c_i,\, \{x_{ji}\}_j\}_i$ and prices $\{p_i,\, w_i\}$ such that*

1. *The labor supply at any industry i is $h_i = \left[(1-\alpha_i)\,\eta \tilde{f}_i\right]^{\frac{1}{1+\varepsilon}}$;*

2. *For any industry i, given the steady state output y_i, the consumption of product i, c_i, and sector j's use of product i as input, x_{ji}, are given by*

$$\begin{aligned} c_i &= \frac{\theta_i}{\tilde{f}_i} y_i, \\ x_{ji} &= \beta \alpha_j \omega_{ji} \frac{\tilde{f}_j}{\tilde{f}_i} y_i; \end{aligned}$$

3. *The set of steady state outputs $\{y_i\}$ is the unique solution to the following*

system

$$y_i = z_i h_i^{1-\alpha_i} \prod_j \left(\beta \alpha_i \omega_{ij} \frac{\tilde{f}_i}{\tilde{f}_j} y_j \right)^{\alpha_i \omega_{ij}}, \forall i.$$

4. The relationship between wages and between prices satisfies

$$\begin{aligned} \frac{w_i}{w_j} &= \frac{h_j}{h_i} \frac{(1-\alpha_i)\tilde{f}_i}{(1-\alpha_j)\tilde{f}_j}, \\ \frac{p_i}{p_j} &= \frac{y_j}{y_i} \frac{\tilde{f}_i}{\tilde{f}_j} \end{aligned}$$

1.3.2 News about Product-Specific Demand As the Driving Force

In the introductory example of motor vehicle production, the series of indices for Vehicle Buying Conditions can be viewed as a measure of the consumers' expected buying capacity of household motor vehicles. Under the lens of this model, this buying capacity of a particular good corresponds to product-specific demand, which is captured by the expectation of the equilibrium consumption share of this good. Therefore, in this subsection, I model the external signals as shocks that can change the expectation of future consumption shares.

Recall the recursive formation of the forecasts

$$\begin{aligned} \tilde{\mathbf{f}}_t &= E_t(\theta_{t+1}) + \beta \Omega' \mathbf{A} E_t\left(\tilde{\mathbf{f}}_{t+1}\right) \\ &= E_t(\theta_{t+1}) + \beta \Omega' \mathbf{A} E_t(\theta_{t+2}) + \beta^2 (\Omega' \mathbf{A})^2 E_t(\theta_{t+3}) + ... \end{aligned} \tag{1.22}$$

hence for any sector i, the expected sales-value added ratio $\tilde{f}_{it}$ summarizes the expectations of all future consumption distribution, $\{\theta_{t+\tau}\}_{\tau=1}^{\infty}$. The weights depend on the input-output structure and the discount factor. To understand the recursive structure of expectations, consider the decision of a specific sector i whose $\tilde{f}_{it}$ is given by

$$\tilde{f}_{it} = E_t\left(\theta_{i,t+1}\right) + \beta \sum_j \alpha_j \omega_{ji} E_t\left(\theta_{j,t+2}\right) + \beta^2 \sum_j \sum_{j'} \alpha_j \omega_{ji} \alpha_{j'} \omega_{j'j} E_t\left(\theta_{j',t+3}\right) + ...$$

where the first term $E_t\left(\theta_{i,t+1}\right)$ is the one-period-ahead forecast of sector i's own consumption share; the second term $\beta \sum_j \alpha_j \omega_{ji} E_t\left(\theta_{j,t+2}\right)$ is the time-discounted weighted sum of i's customers' two-period-ahead forecasts, weighted by the importance of product i in the production of customer j's output; the third term $\beta^2 \sum_j \sum_{j'} \alpha_j \omega_{ji} \alpha_{j'} \omega_{j'j} E_t\left(\theta_{j',t+3}\right)$ is the discounted weighted sum of three-period-ahead forecasts of i's customers' customers, twice weighted; and so on for further terms.

To formalize the process of updating beliefs while preserving the tractability of the model, consider a specific form of the set of external signals M_t received at time t such that M_t contains information about future product-specific consumption demand that arrives T periods ahead, $T \geq 1$. Specifically,

$$M_t = \left\{\mathbf{m}_{t+1}^t, \mathbf{m}_{t+2}^t, ..., \mathbf{m}_{t+T}^t\right\}$$

such that for each $\tau = 1, ..., T$, $\mathbf{m}_{t+\tau}^t$ is drawn independently from a distribution determined by $\theta_{t+\tau}$,

$$\mathbf{m}_{t+\tau}^t \sim \Xi^m\left(\cdot|\theta_{t+\tau}\right)$$

All sectors receive the same signals in each period and update their expectations of $\theta_{t+\tau}$.

The structure of the signal set is designed to capture the idea that agents receive information and form expectations of future demand. The information is allowed to accumulate over time and the forecasts get more precise as available information grows. For example, comparing the current forecasts of demand for cars in one year and demand for cars in five years, one should expect the former to be more reliable because there are more signals observed. The assumption of multinomial signals is for technical simplicity. It accompanies the common prior Dirichlet distribution of product-specific preference vector $\theta_{t+\tau}$ so that the updated posterior of the preference vector remains a Dirichlet distribution, which allows for simple explicit expression of the updated expectations. I do not wish to over-emphasize this functional form assumption on the signals for it merely complements the Dirichlet distribution in a Bayesian updating process.

At each time t, $\mathcal{I}_t$ contains $T-\tau+1$ signals of $\theta_{t+\tau}$, $1 \leq \tau \leq T$, which are $\mathbf{m}_{t+\tau}^{t-(T-\tau)}$, $\mathbf{m}_{t+\tau}^{t-(T-\tau)+1}$, ..., and $\mathbf{m}_{t+\tau}^{t}$. For example, $\mathbf{m}_{t+T}^{t}$ is the only signal vector of θ_{t+T} at time t, and there are T signal vectors of θ_{t+1}, of which the first one was received $T-1$ periods earlier. In fact, the explicit forms of the posterior expectations are given by:

$$\begin{aligned}
E_t\left(\theta_{t+1}\right) &= E\left(\theta_{t+1}|\mathbf{m}_{t+1}^{t-(T-1)}, \mathbf{m}_{t+1}^{t-(T-2)}, ..., \mathbf{m}_{t+1}^{t}\right) \\
E_t\left(\theta_{t+2}\right) &= E\left(\theta_{t+2}|\mathbf{m}_{t+2}^{t-(T-2)}, \mathbf{m}_{t+2}^{t-(T-3)}, ..., \mathbf{m}_{t+2}^{t}\right) \\
&\vdots \\
E_t\left(\theta_{t+T}\right) &= E\left(\theta_{t+T}|\mathbf{m}_{t+T}^{t}\right).
\end{aligned}$$

The longer the time horizon, the less precise the available information is. $\mathcal{I}_t$ does not contain any additional information on future periods beyond $t+T$ besides the prior distribution of consumption shares. Therefore,

$$\tilde{\mathbf{f}}_t = \sum_{\tau=1}^{T} \beta^{\tau-1} \left(\Omega'\mathbf{A}\right)^{\tau-1} E_t\left(\theta_{t+\tau}\right) + \beta^T \left(\Omega'\mathbf{A}\right)^T \tilde{\mathbf{f}}.$$

I call the changes in the expectations $E_t\left(\theta_{t+\tau}\right)$ the result of forecast shocks, which in turn affect the forecasts and decisions of firms. Under the assumption of a common information set, I am able to maintain tractability of the model and step aside from the complexity of extracting information from prices and higher order beliefs when agents have heterogeneous information. The forecast shocks have the following features. First, he shocks to expectations have real impact because agents are forward-looking in making decisions. In particular, the forecasts enter the input purchase decisions of firms due to the timing restriction that requres the firms to decide on the amount of inputs without knowing future prices. In principle, the news on future productivity should affect the current decisions via the same intertemporal-concern channel. However, under the functional assumptions of this model, the expected change in productivity and the expected change in price cancel each other out. Therefore, I can isolate the effect of the novel news-on-demand shocks. Second, the shocks are transmitted upwards from customers to suppliers through the input-output links and these are the only upward-transmitting shocks under the model specifications. In reality, shock transmission in the economy is more plausibly bi-directional than either downwards only or upwards only and this should be the case for transmission of productivity shocks, news shocks, or other shocks. In the static variant of this model studied by Acemoglu et al. (2012), the productivity shocks have immediate impact on real output of both upstream and downstream sectors because of market clearing prices. Third, the shocks affect the size distribution of the sectors and have decaying and lasting real effects over time. Compared to productivity shocks and news shocks to productivity, which cannot generate changes in the distribution of sectors' sales shares, the effects of forecast shocks on prices and on quantities do not cancel each other out. Hence they show in sales, consumption expenditure, input expenditure, etc. Moreover, the effects on quantities last and decay over time through the firms' input purchase decisions. Lastly, in addition to its more conventional role in prolonging and propagating real effects on outputs, the input-output structure plays an essential role in determining the scale and direction of a forecast shock. Suppose a signal in favor of sector j' arrives at time t such that all agents expect that the consumer will spend relatively more on good j' at time $t+\tau$, hence $E_t\left(\theta_{j',t+\tau}\right)$ goes up by $\Delta_{j'}$ (and surely the expectations of consumption shares on other goods will decrease accordingly). The scale and direction of the impact of this change depends both on the position of sector j' in the production network and on how large τ is. Notice that the influence of the change $\Delta_{j'}$ varies as a result of two simultaneous effects as τ becomes larger: (1) the change will affect more sectors through the input-output connection in the upstream direction, while (2) the weight on $\Delta_{j'}$ gets smaller and more heavily discounted. Furthermore, the shocks to expectations of the future, even the distant future, have prolonged real effects through the input purchase decisions. The net effect of such a change is further discussed and demonstrated quantitatively in the simulation section.

Unlike the demand parameter η_t and the productivity indices $\left\{z_{it}\right\}_i$, both of which are conventional in the business cycle literature and directly affect the fundamentals, the forecast shocks and the interplay between forecast shocks and input-output connections are newly introduced in this paper. Moreover, they are not explicitly related to any of the *current* fundamental variables in the economy, but rather reflect the future economic conditions. The formation of forecasts also captures the second role the supplier-customer network plays in reality: it allows communication of information between a supplier and a customer when trading. In a very stylized fashion, the forecast formation process shows how a firm or an industry gathers and exploits information from its business activity, and how it makes production and input purchase decisions based on this information.

Notice that external signals affect the economic activities only by changing agents' forecasts, and the change of forecasts is driven by the arrival of external signals only. Consequently, one can understand how the arrival of external signals affects the economy by studying how the changes in forecasts affect the economy. In the following analysis, I show that the sector specific forecast $\tilde{f}_{it}$ has a non-trivial effect both on the distribution of sales and on total output. Not only does $\tilde{f}_{it}$ affect industry i, but it also has a direct impact on other industries that are connected to i through the input-output link. Sector i's output is increasing in its own forecast $\tilde{f}_{it}$. Also, sector j's output is increasing in i's forecast if sector i uses product j as an input. In other words, the impact of one sector's forecast on the output level of another sector goes upstream from a customer (i) of a certain product (j) to its producer (j).

$$\begin{aligned} \frac{dy_{it}}{d\tilde{f}_{it}} &> 0 \\ \text{and } \frac{dy_{jt}}{d\tilde{f}_{it}} &> 0 \text{ if } \alpha_i \omega_{ij} > 0 \end{aligned}$$

Sector i increases its input expenditure when its own forecast increases. Sector j decreases its output expenditure when i's forecast increases.

$$\begin{aligned} \frac{d\left(P_{it}^{X}X_{it}\right)}{d\tilde{f}_{it}} &= \beta\alpha_i P_t^C C_t\left(1-\beta\alpha_i P_t^C C_t\right) > 0 \\ \text{and } \frac{d\left(P_{jt}^{X}X_{jt}\right)}{d\tilde{f}_{it}} &< 0 \text{ if } \alpha_j > 0 \end{aligned}$$

When sector i's forecast increases, whether its share of total industrial sales increases or not depends on how heavily sector i uses its own product as an input. In fact, most of the sectors retain a large fraction of their outputs for each period.

$$\frac{ds_{it}}{d\tilde{f}_{it}} = \beta\alpha_i P_t^C C_t\left(\omega_{ii} - s_{it}\right)\begin{cases} \geq 0 \text{ if } \omega_{ii} \geq s_{it} \\ < 0 \text{ if } \omega_{ii} < s_{it} \end{cases}$$

Similarly, in response to the same increase in i's forecast, sector j's share of sales increases only if i is an important customer of product j, that is, out of i's input purchases, the fraction spent on product j is larger than sector j's share of total industrial sales.

$$\frac{ds_{jt}}{d\tilde{f}_{it}} = \beta\alpha_i P_t^C C_t\left(\omega_{ij} - s_{jt}\right)\begin{cases} \geq 0 \text{ if } \omega_{ij} \geq s_{jt} \\ < 0 \text{ if } \omega_{ij} < s_{jt}. \end{cases}$$

Hence, the change in distribution of shares of industrial sales reflects the impact of changing forecasts. Such change travels upstream through supplier-customer connections similar to the way the impact of forecasts on output levels does. However, unlike the case of output levels, whether and how much the sales share will increase depends on the importance of the seller's product as an input to the customer.

1.3.3 Impact Paths of News

As in previous discussion, due to the forward-looking and recursive fashion of the forecast formation, a change in expected future forecasts shows up immediately in current forecasts via the supplier-customer connections, which in turn affect the distribution of sales shares in the present period. The real effect of such change will last into the following periods because of the dynamic input choices. In this section, I illustrate the model mechanism by a counterfactual exercise. The parameterization of the model is discussed in the calibration and estimation section.

To highlight the effects of changes in forecasts, let the realization of the product-specific preference vector, i.e., the consumption shares, be fixed at the mean, $\theta_t = \theta$, while preserving the common prior distribution of θ_t. Moreover, fix the level of productivity $\mathbf{z}_t = \mathbf{z}$ and fix the aggregate preference $\eta_t = \eta$. Consider a simplified version of the posterior updating process, in which only one external signal arrives so the forecasts are updated once upon its arrival. Let t_0 be the time of impact when an external signal arrives. Before this period, the agents form stationary forecasts $\tilde{\mathbf{f}}$ and expect no change in future consumption shares. Let the signal be such that the updated expected consumption share of good i in T periods changes relatively by fraction δ, and the other shares change accordingly:

$$\begin{aligned} E_{t_0}\left(\theta_{i,t_0+T}\right) &= \frac{\theta_i\left(1+\delta\right)}{1+\delta\theta_i}, \\ E_{t_0}\left(\theta_{j,t_0+T}\right) &= \frac{\theta_j}{1+\delta\theta_i}, \forall j \neq i. \end{aligned}$$

Here T is called the target time and i the target sector. In addition, since there is no further signal in the following periods, $E_t\left(\theta_T\right) = E_{t_0}\left(\theta_T\right)$ for $t_0 \leq t \leq T-1$.

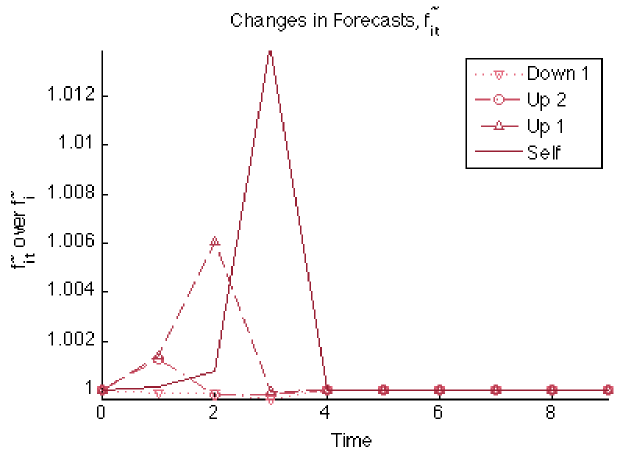

Figure 1.5: Response of forecasts $\tilde{f}_{jt}$ on the selected chain. X axis is time and Y axis is the ratio between forecasts $\tilde{f}_{jt}$ and stationary forecasts $\tilde{f}_j$. Solid line: underlining sector 324; dashed line: direct upstream sector 3; broken line: 2nd-order upstream sector 532; dotted line: immediate downstream sector 481. Impact time $t_0 = 1$, $i=$ NAICS 324, change in θ_i: $\delta = 0.1$, $T = 3$. All prior uncertainty resolves at $t_0 + T = 4$.

For the purpose of illustration, consider a specific experiment in which $\delta = 0.1$, $T = 3$, and sector i that undertakes the δ change is $i =$ 'Petroleum and coal products' with NAICS code 324. In general, at the industry level, an industry is connected to several others by the input-output relation. Hence, instead of production chains, one observes a production network measured by the matrices $\mathbf{A}$ and Ω. I pick a "chain-like" subset of this network for better demonstration and the logic holds for more general cases. One of Petroleum and coal products's major downstream sectors, namely its customers, is NAICS 481'Air transportation'. Its most important immediate supplier is NAICS 3 'Oil and gas extraction' whose major suppliers include NAICS 532RL 'Rental and leasing services and lessors of intangible assets'.

Now, at time $t_0 = 1$, the economy receives the external signal and all agents update their expected future consumption shares: The expected consumption share of i at time $t_0+T=4$ increases about 10% and the expected shares of other sectors adjust accordingly. The updated expectations show up directly in the forecasts upon impact. This impact then affects the distribution of shares of industrial sales, and the real terms: output levels, consumption, input purchase, etc. Figure 1.5 shows the changes of forecasts formed by sectors on the "chain" over time. All four sequences of forecasts change immediately upon the arrival of the signal at time 1. However, each spikes at a separate time. For the underlining sector, the forecast sequence spikes at time 3, one period before the uncertainty is resolved. For its supplier and the supplier's supplier, the maximum forecasts occur at time 2 and time 1, respectively. The customer's forecast does not change significantly. It is important to notice how the impact of shocks to expectation is transferred in the upstream direction. Upon impact, the underlining sector anticipates a higher consumption share of its own product in three periods and its forecasts of future sales relative to total consumption are adjusted accordingly. Its direct supplier anticipates the same and the increased expectation of customer sales will drive the purchase of the supplier's product as input, which will happen in two periods. The same logic explains the spike in the forecast of the supplier's supplier. Therefore, a shock to the forecasts acts as a demand-side shock and travels upstream through the supplier-customer connections. In the downstream direction, however, the main blow of the forecast shock does not directly affect the customer. Once the uncertainty resolves, in this case when the signal is proven to be "wrong", the forecasts instantly adjust back to the stationary levels.

The responses of the employed labor, real output, consumption, and input uses

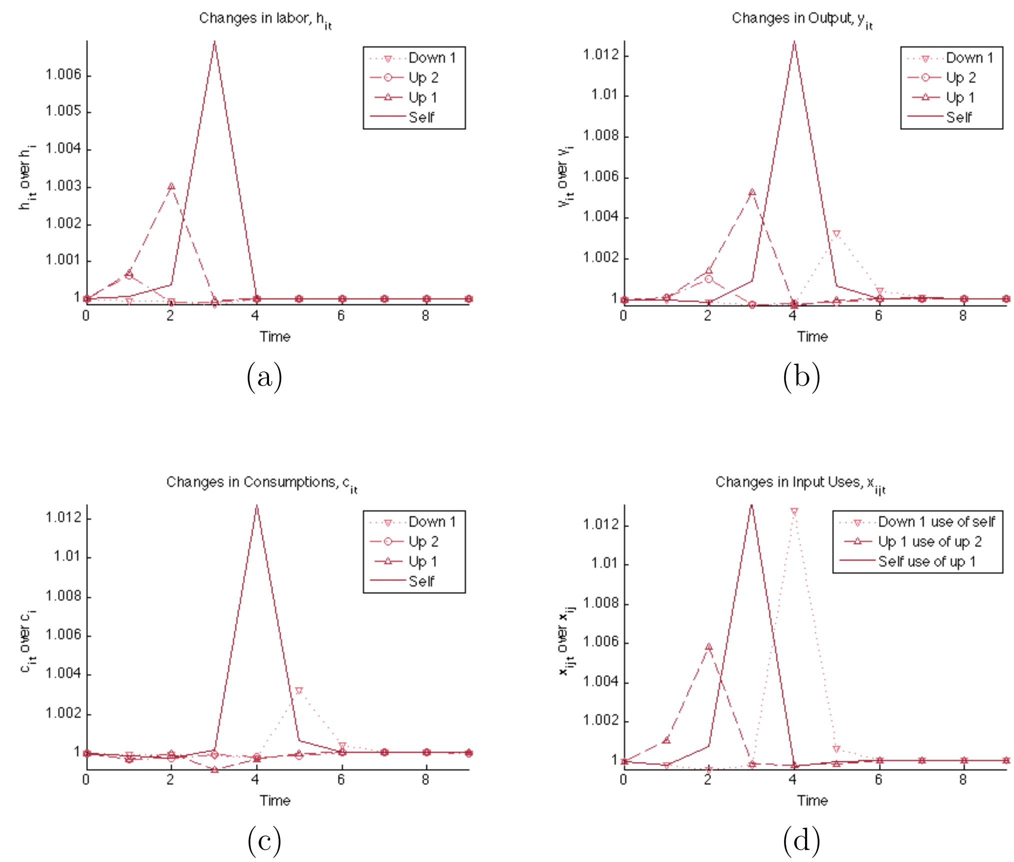

Figure 1.6: Response of (a) labor, (b) output, (c) consumption, and (d) input uses on the selected chain. X axis is time and Y axis is the ratio between variable value in each period and corresponding stationary value. Industry names are in the caption of Figure 1.5.

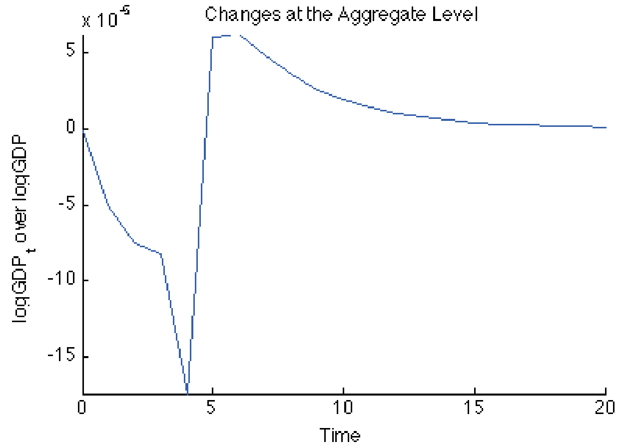

Figure 1.7: Response of aggregate value added. X axis is time and Y axis is the ratio between $\log(GDP_t)$ and steady state $\log(GDP)$.

are shown in panels (a) to (d) in Figure 1.6, respectively. Obviously, not only does the transitory shock to forecasts have real effects but these real effects also last because of the one-period-ahead input decisions. At each time, the real effects come from two sources: one is the change in the forecasts formed in that period and the other is the lasting impact of changes in previous periods through the input decisions. Since the forecasts return to the stationary levels at time 4, deviation of output, consumption, and input uses from their steady state levels from time 4 onwards is the cumulated result of the effect from the latter source. This is the same as the lasting impact of a transitory productivity shock.

At the aggregate level, Figure 1.7 shows how the logarithm of real GDP, denoted as $\log(GDP_t)$, responds to the same shock to forecasts, where $\log(GDP_t) = \log(C_t)$. With normalized prices, changes in real GDP are in fact changes in the consumption index and the logarithm of real GDP is equivalent to the utility gained from consuming the basket of goods. The impact of the shock to forecasts on $\log(GDP_t)$ is, in general, not monotonic over time. Similar to the case of consumption of an individual product, before the uncertainty resolves, $\log(GDP_t)$ is under the direct impact of current-period forecasts and the indirect impact from previous forecasts carried through the input-output links; while after the uncertainty resolves, the lasting effect is solely due to the inputs.

This section demonstrates the the mechanism of the transmission of news, using an arbitrary and fixed news process. The quantitative analysis in the following section takes further steps in examining the effects of both preference shocks and news shocks with the estimated driving forces.

1.4 Quantitative Explorations

An important feature of the model is that in equilibrium, the evolvement of the forecasts sequence $\left\{\tilde{\mathbf{f}}_t\right\}$ relies only on the distributions of future product-specific preference shocks $\{\theta_\tau\}_{\tau>t}$ and the news about future demand. Once the values of forecasts are given, the rest of the equilibrium components are also determined. The parameterization of the model goes backwards: I calibrate part of the model to match the observed variables in the data and then I estimate the distribution of the news. Using the estimation results, I show the importance of the news about future demand as a driving force of the aggregate economic volatility.

1.4.1 Calibration

The main data set used to calibrate and estimate the model is taken from the Annual Industry Accounts at the Bureau of Economic Analysis, which contains information on the annual production and input usage of 65 private industries from

Table 1.1: Parameter Values

Predetermined Parameters	
$\beta = 0.96$	Discount factor
$\varepsilon = 1$	Disutility in labor
$z = 10$	Productivity
Calibrated Parameters	
$\eta = 9.50$	Mean aggregate demand s.t. $\bar{h}_i = 1/3$
$\mathbf{A}$	Input expenditure shares
Ω	Input requirements
θ	Mean product-specific demand

1997 and 2012.

In particular, the annual use and make tables set the targets of calibrating the most important set of parameters, which contains the shares of input goods in each industry, $\mathbf{A} = diag\,(\alpha_1, ..., \alpha_n)$; the specific share of each certain good used as input, Ω with $\Omega\,(i, j) = \omega_{ij}$; and the mean of the prior distribution of product-specific demand parameters, $\theta = (\theta_1, ..., \theta_n)$. The use table of each year is essentially a matrix that shows the uses of commodities by industries as intermediate inputs and by final users from 1997 to 2012, while the make table shows the quantity of each commodity produced by the industries. Both the uses and makes are measured in terms of dollar expenditures.

The model assumes that the production technologies determined by $\mathbf{A}$ and Ω are constant over time, hence the values are chosen to match the make and use tables of year 2007. Specifically, at the steady state, each element α_i of $\mathbf{A}$, $i = 1, ..., n$, is proportional to the fraction of total industrial sales of industry i used to purchase input goods from all industries.

$$\alpha_i = \frac{\text{input expenditure by } i}{\beta \times \text{total sales of } i}.$$

Each row i of matrix Ω determines industry i's use of all industries' products as inputs and every element ω_{ij} is the fraction of i's input expenditure on j's output.

$$\omega_{ij} = \frac{i\text{'s purchase of } j\text{'s output}}{\text{input expenditure by } i}.$$

In equilibrium, the realized product-specific preference parameters θ_t are the current-period total consumption expenditure shares of each product. Hence θ_t can be chosen such that the equilibrium composition of consumption expenditure lines up perfectly to the data. The consumption expenditure on each commodity is measured as the final use of that commodity, including the amount imported.

$$\theta_{it} = \frac{\text{Final use on } i \text{ at } t}{\text{Total final uses at } t},$$

and the unconditional mean of the consumption shares θ is set at the simple average over the 1997 to 2012 period,

$$\theta = \frac{\sum_t (\text{Final use of each commodity at } t)}{\sum_t (\text{Total final uses at } t)}.$$

The discount factor β is set to be 0.96 to match the annual frequency of the data set. In addition, I set $\varepsilon = 1$ so that the disutility in labor has a quadratic form. The unconditional mean of the aggregate preference shock, η, is set at 9.49 which ensures the average labor supply at the steady state is 1/3. Moreover, since the process of aggregate preference η_t and that of the sector-specific productivity $\{z_{it}\}_{i=1}^{n}$ do not affect the sequence of $\left\{\tilde{f}_{it}\right\}$ under the assumption of the model, the uncertainty is shut down such that the aggregate preference is fixed over time, $\eta_t = \eta$, $\forall t$ and all sectors share the same time-invariant productivity $z_{it} = z_i = z$, $\forall i, t$. The the productivity z sets the scale of the model solution and is set at 10.

1.4.2 Estimation

The analysis of the forecast formation does not require any specific distributions of the product-specific preference shocks or of the news process. For the purpose of estimation and simulation, I impose further assumptions on these distributions, $\Xi^{\theta}(\cdot)$ and $\Xi^{m}(\cdot|\theta_t)$.

Assume that the independent and identical prior $\Xi^{\theta}(\cdot)$ of the product-specific prefence shocks θ_t is a Dirichlet distribution of order n with concentration parameters θ and a scale parameter $\kappa > 0$,

$$\theta_t \sim Dirichlet_n(\kappa\theta)$$

hence $E(\theta_t) = \theta$ and κ determines the scale of the variance-covariance matrix of θ_t such that

$$\begin{aligned}
\left(\sigma_i^{\theta}\right)^2 &= var(\theta_{it}) = \frac{\theta_i(1-\theta_i)}{\kappa+1}, \forall i, \\
\sigma_{ij}^{\theta} &= cov(\theta_{it}, \theta_{jt}) = -\frac{\theta_i\theta_j}{\kappa+1}, \forall i \neq j.
\end{aligned}$$

The unconditional expectation θ is pinned down in the calibration. Note that the correlation coefficient between any pair of $(\theta_{it}, \theta_{jt})$ is always negative and has the value of $-\sqrt{\frac{\theta_i \theta_j}{(1-\theta_i)(1-\theta_j)}}$, $\forall i, j$, regardless of the value of κ. To accommodate the Dirichlet prior, let the news that contains the information on θ_t be drawn from a multinomial distribution. Specifically, for each t, the distribution of news $\mathbf{m}_t^{t-\tau}$ about θ_t that arrives in period $t-\tau$ is

$$\mathbf{m}_t^{t-\tau} \sim Multinomial\left(N, \theta_t\right), \forall \tau = 1, ..., T,$$

where the integer $N \geq 1$ is the number of trials and θ_t represents the probability associated with each possible outcome in one trial. Conditional on the news, the posterior distribution of θ_t remains Dirichlet. In each period t, the posterior distributions of future $\theta_{t+\tau}$ have the following expectations,

$$\begin{aligned} E_t\left(\theta_{t+1}\right) &= \frac{\kappa\theta + \sum_{\tau'=0}^{T-1} \mathbf{m}_{t+1}^{t-\tau'}}{\kappa + TN} \\ E_t\left(\theta_{t+2}\right) &= \frac{\kappa\theta + \sum_{\tau'=0}^{T-2} \mathbf{m}_{t+2}^{t-\tau'}}{\kappa + (T-1)N} \\ &\vdots \\ E_t\left(\theta_{t+T}\right) &= \frac{\kappa\theta + \mathbf{m}_{t+T}^{t}}{\kappa + N}, \end{aligned}$$

hence

$$\tilde{\mathbf{f}}_t = \sum_{\tau=1}^{T} \beta^{\tau-1} \left(\Omega'\mathbf{A}\right)^{\tau-1} \frac{\kappa\theta + \sum_{\tau'=0}^{T-\tau} \mathbf{m}_{t+\tau}^{t-\tau'}}{\kappa + (T+1-\tau)N} + \beta^T \left(\Omega'\mathbf{A}\right)^T \tilde{\mathbf{f}}.$$

The specific information structure discussed in the theoretical model is very stylized, according to which the changes in forecasts solely come from the news about future product-specific demands. The remaining parameters to be estimated are κ, N, and T, where κ controls the scale of the variance matrix of the prior distribution of θ_t, N determines the precision of news, and T sets how far into the

Table 1.2: Forecast Process Parameters

Parameters		Estimated Values
κ	Scale of prior variance of θ_t	15300
N	Precision of news	9100
T	Horizon of news	1

future the news can reach.

The estimation consists of two steps. In the first step, κ is picked to minimize the distance between the model variance of θ_t and the data counterpart, namely, $var(\theta_{it})$, $\forall i = 1, ..., 65$. In the second step, (N, T) is jointly determined with κ given. Observe that, by Proposition 1, the forecasts enter linearly the expressions for total sales, input expenditures, and value added of industries divided by aggregate value added. Hence, the variation in these variables reflects that in the forecasts, which identifies N and T. Then (N, T) is picked to minimize the difference between the model and the data variances, that is, $var(\frac{s_{it}}{P_t^C C_t})$, $var(\frac{P_{it}^X X_{it}}{P_t^C C_t})$, and $var(\frac{s_{it}-P_{it}^X X_{it}}{P_t^C C_t})$, $\forall i = 1, ..., 65$. The detailed description of the estimation procedure is in the appendix.

According to the estimation result, at the annual frequency, the news that arrives in each period contains only information on the product-specific preference in the following period, $T = 1$. Consequently, in the model language, given the estimated (κ, N)

$$\tilde{\mathbf{f}}_t = \frac{\kappa\theta + \mathbf{m}_{t+1}^t}{\kappa + N} + \beta\Omega'\mathbf{A}\tilde{\mathbf{f}}.$$

The limited horizon of the news $(T = 1)$ is partially due to the relatively low frequency of the annual data and the modeling assumption that news updates once in each period.

1.4.3 Quantitative Results

In this section, I examine the quantitative performance of the model and look at the aggregate volatility, industry-level volatility, and industry comovement. Additionally shown, the positive relationship between input prices and input uses observed in the data can be explained by the news transmission mechanism. Last but definitely not least, in the absence of productivity shocks, the model is capable of generating sizable volatility in the measured Solow residuals, both at the aggregate level and at the industry-level.

Specifically, I do the following three experiments, each of which starts from the steady state of the model and evolves under a particular driving force. In the first experiment (E1), the product-specific preference in each period, θ_t, is independently drawn from the Dirichlet distribution with concentration parameters $\kappa\theta$, the news is informative, $\mathbf{m}_{t+1}^t \sim MN(N, \theta_{t+1})$, and the prior distribution of θ_t is "correct". The θ_t sequence in the second experiment (E2) is fixed at its mean, $\theta_t = \theta$, the news is then iid, $\mathbf{m}_{t+1}^t \sim MN(N, \theta)$, while the agents holds the Dirichlet prior. In the third experiment (E3), the only driving force is the changing θ_t. In each period of each simulation, the ideal price index P_t^C is normalized to 1.

1.4.3.1 Aggregate Volatility

The aggregate volatility is measured as the standard deviation in the growth rate of real GDP generated by the private sector. In fact, aggregate GDP is accounted as the sum of value added across all industries and, equivalently, as the sum of total final consumptions. During the sample period, between 1997 and 2012, the standard deviation of private GDP growth rate is 2.18%. The model counterpart

Table 1.3: Volatility in Aggregate GDP

	Data 1997-2012	E1: Both	E2: News	E3: Preference
Std.Dev.(ΔGDP)	2.18	0.90	0.08	0.90
Model/Data	100	41.28	3.67	41.28

is the standard deviation in the growth rate of total consumption index, denoted as ΔV_t

$$\Delta V_t = \log \frac{C_t}{C_{t-1}}.$$

Note that V_t is also the aggregate value added. Table 1.3 reports the simulation results.

The driving force of the economy in E1 is a mixture of product-specific preference shocks θ_t and news $\mathbf{m}_{t+1}^t$ that reflects the changing θ_t, which is capable of generating the aggregate volatility 0.90%, which is around 41% percent of the actual volatility in the growth rate of real GDP. In E2, the news process $\mathbf{m}_{t+1}^t$ is the only source of shocks. Note that because θ_t is fixed at θ, the distribution of news is fixed as well. With all relevant fundamentals fixed at their steady state levels, the iid news alone generates small fluctuation at the aggregate level, about 3.67% of the scale of actual fluctuation. E3 illustrates the importance of the product-specific preference shocks θ_t, without news arrivals, capable of generating aggregate fluctuation at a scale very similar to E1.

1.4.3.2 Industry-level Volatility

Now, we zoom in and take a closer look at each of the experiments compared with the data. For each industry, I compute the time-series standard deviations in

Table 1.4: Average Industry-level Volatility

		Data 1997-2012	E1: Both	E2: News	E3: Preference
(1)	$\overline{\sigma^{\Delta s/V}}$	0.15	0.12	0.008	0.12
(2)	$\overline{\sigma^{\Delta v/V}}$	0.081	0.13	0.024	0.12
(3)	$\overline{\sigma^{\Delta u/V}}$	0.13	0.033	0.027	0
(4)	$\overline{\sigma^{\Delta s}}$	5.86	6.11	0.33	6.13
(5)	$\overline{\sigma^{\Delta v}}$	8.04	15.85	2.81	14.10
(6)	$\overline{\sigma^{\Delta u}}$	11.23	3.85	2.95	0.90
(7)	$\overline{\sigma^{\Delta u/s}}$	3.20	3.68	1.31	2.84

the year-to-year changes in the following ratios: sales over aggregate value added or GDP ($\Delta s/V$), value added over aggregate value added ($\Delta v/V$), and input expenditure over aggregate value added ($\Delta u/V$), as well as standard deviations in the annual growth rates of real gross output or sales (Δs), real value added (Δv), and real input expenditure (Δu), input expenditure over sales ($\Delta u/s$). Table 1.4 summarizes the volatility in these variables averaged across industries.

Rows (1) to (3) of E1 show part of the estimation results. On average, the estimated parameters capture the volatility in sales to GDP ratio ($\Delta s/V$), while generating too much volatility in the value added contribution by industry ($\Delta v/V$) and not enough in the input expenditure to GDP ratios ($\Delta u/V$). This is due to the structure of the model, where both preference shocks and forecasts affect the sales and value added, while input expenditure depends on forecasts (hence news) only. Consequently, while preference shocks alone can generate relatively large volatility in the sales and value added by industry, the news process contributes the majority of the model-generated volatility in input expenditure. The reason is that, in the model environment, the time variant forecasts drive the changes in an industry's input expenditure decision. Without news arrival, forecasts are

Table 1.5: Industry-level Comovement

		Data 1997-2012	E1: Both	E2: News	E3: Preference
(1)	$\overline{\rho^{\Delta s,\Delta V}}$	0.56	0.14	0.38	0.14
(2)	$\overline{\rho^{\Delta v,\Delta V}}$	0.36	0.04	0.03	0.05
(3)	$\overline{\rho^{\Delta u,\Delta V}}$	0.44	0.31	0.05	1
(4)	$\overline{\rho_{i,j}^{\Delta s}}$	0.33	0.01	0.15	0.01
(5)	$\overline{\rho_{i,j}^{\Delta v}}$	0.13	-0.01	-0.01	-0.01
(6)	$\overline{\rho_{i,j}^{\Delta u}}$	0.22	0.09	-0.01	1

constant over time, and consequently the input expenditure is proportional to the aggregate consumption or aggregate value added, which explains the zero in row (3). Rows (4) to (6) shows similar results to previous rows in terms of growth rates in sales, value added, and input uses, instead of ratios over GDP. Volatility in row (7) is a feature of the model in that the fraction of sales that an industry uses to purchase inputs reflects its forecasts of future sales relative to current sales, hence this fraction exhibits fluctuation over time. In absence of the preference shocks and news process, productivity shocks cannot generate such changes.

1.4.3.3 Industry-level Comovement

The assumption on driving forces of the model, either preference shocks or news process, explicitly imposes negative relationship between industries. Meanwhile, industries in the model economy are interconnected via supplier-customer links, which has the potential of creating positive correlation between industries. It is not unambiguous which force dominates when it comes to industry comovement. Table 1.5 shows the correlation coefficients that can measure the comovement and compares the data and the models. In the data, an industry's economic activity is on average positively correlated to that at the aggregate level and industries also comove with each other, which is the most important feature of business cycles. Rows (1) to (3) are average correlation coefficients between aggregate real GDP growth rate (ΔV) and an industry's sales growth rate (Δs), value added growth rate (Δv), and its input expenditure growth rate (Δu). Although shocks are cross-sectionally negatively correlated and common aggregate shocks are absent, the models can generate positive correlations between an industry and the whole economy, with various scales. Rows (4) to (6) show the correlations of growth rates of sales (Δs), value added (Δv), and input uses (Δu) between a pair of industries, averaged across 65×64 pairs. The effect of cross-sectionally negatively correlated shocks dominates in the pair-wise correlation of industry value added, while the effect of supplier-customer links between industries dominates in the sales and input uses. Note that, in E3, an industry's input expenditure is always a fixed proportion of the aggregate GDP, therefore we see the perfect correlation in rows (3) and (6).

1.4.3.4 Input Prices and Uses

Another interesting issue is the relationship between the input prices and the input uses. Each industry needs to buy a bundle of input goods facing the input prices, which form an industry-specific input price index. The data show that, on average, the growth rates of input price index faced by an industry (Δp^{Input}) and the growth rates of this industry's input quantity (Δx) are positively correlated. In addition, the input price index and the changes in the industry's fraction of sales used towards input purchases ($\Delta u/s$) are positively correlated as well. Results in Table 1.6 show that the model with news process as the only driving force, E2, is

Table 1.6: Input Prices and Uses

	Data 1997-2012	E1: Both	E2: News	E3: Preference
$\rho\left(\Delta x, \Delta p^{Input}\right)$	0.24	-0.08	0.24	-0.60
$\rho\left(\Delta u/s, \Delta p^{Input}\right)$	0.34	-0.14	0.29	-0.27

capable of capturing both of the positive correlations.

Intuitively, an industry expecting better sales condition in the following period is willing to spend more on its inputs, which shows as both higher level and higher shares of input expenditure, and which drives up the input price index. When there are preference shocks only, an increase in the final demand for a certain product drives up its price, and because more of this product is consumed by the household, less is used as an input purchased by the industries, which decreases the input quantities of these industries. Note that, if the productivity shocks act as the pure driving force, the ratio of an industry i's input expenditure over its sales remains constant at $\beta\alpha_i$.

1.4.3.5 Volatility in Measured Solow Residuals

Productivity shocks are omitted from the three experiments, however, using the Solow residuals as the measure of the productivity shocks, volatility can still be observed. In fact, this exercise provides a possible explanation for the volatility in measured productivity in reality.

To illustrate, consider each experiment as the true economy and treat the generated data as the observed variables. Then, suppose we view this economy through the lens of a standard neoclassical model and calculate the Solow residuals from the value added or GDP. To be consistent with the assumption that labor is the

Table 1.7: Volatility in Aggregate Solow Residuals

	Data 1998-2012	E1: Both	E2: News	E3: Preference
$\sigma^{\Delta a}$	1.08	1.01	0.08	1.01

only value added input, consider a model with a constant returns to scale aggregate production technology for the value added V_t,

$$V_t = a_t H_t, \tag{1.23}$$

where H_t is the total labor input at time t, $H_t = \sum_i h_{it}$, and a_t captures the measured Solow residuals, which in this case is no different from the labor productivity,

$$\log(a_t) = \log(V_t) - \log(H_t). \tag{1.24}$$

The volatility measure is the standard deviation in the growth rate of a_t, denoted as $\Delta \log(a_t) = \log(a_t/a_{t-1})$. Table 1.7 contains the results.

In the absence of productivity shocks, each experiment exhibits a significant amount of fluctuation in the growth rate of aggregate labor productivity comparable to the data counterpart. This mechanism sheds light on the puzzling observation that the measured productivity fluctuates significantly over time. While it is conventional to interpret the changes in productivity as evolvement in production technology, negative productivity growth does not seem to be convincingly justified as a slide back. Therefore, it is important to explore other explanations. Table 1.7 shows that either shifting the household's preference or receiving news about the future preference or both can potentially explain the volatility in the measured productivity. The preference shocks in this model change the relative preference for all products that the representative household consumes, therefore the relative prices of the products are changed. The labor supplied to each industry is also affected. Consequently the contribution to the aggregate value added by each industry changes as well, which is the main channel through which the preference shocks generate volatility in the measured productivity. The pure preference shocks have the direct effect on industry i only if there is a shock to θ_{it}, and the indirect effect takes place due to the input requirement through price changes. The news takes effect through the forecasts, which alter the input expenditure shares and also the labor supply. It differs from the preference shocks in that when news changes forecast of industry i in period t, $\tilde{f}_{it}$, it directly affects the input expenditure made by industry i and, meanwhile, it affects the sales and labor requirement of all the supplier industries of i. The effect then trickles down in the following periods along the supplier-customer links. Note that, in the experiment E2, the preferences are fixed over time and therefore the news are drawn from a fixed distribution, which results in a relatively small yet non negligible volatility in the measured productivity compared to the other two experiments.

1.5 Concluding Remarks

The paper develops a dynamic multi-sector production network model in which firms receive external information on the future demand structure. Since firms are connected via an input-output network, news on the future demand of an individual industry has a global effect. Shocks to future forecasts are transferred upwards through the supplier-customer connections in the network, from the buyer of an input good to the producer. The updated forecasts are reflected in the firms' decision on input expenditure, as well as the suppliers' sales, labor input, and the value added. The effect is shown both at the industry level and at the aggregate level. The model is designed to capture the asymmetry in the transmission of individual shocks in the network, especially the customer-to-supplier direction which cannot be explained by the conventional productivity shocks. The quantitative results demonstrate the model's capability of generating the economic volatility at the aggregate level and at the industry level. The news about future demand can also explain the positive relationship between input prices and input uses observed in the data. Perhaps more importantly, the model points out a potential explanation to the volatility in the measured productivity.

There are interesting issues worth addressing in future research and I briefly discuss some of them here. In terms of modeling, it may be fruitful to consider capital accumulation and inventory management. In this paper, for simplicity, I treat all the intermediate inputs equally and assume full depreciation after one period, and I assume that labor is the only value-added input. In fact, intermediate inputs have different rates of depreciation, which is how the BEA draws the line between capital goods and other materials. My conjecture is that considering the flows of capital goods between industries and allowing capital to accumulate over time may (1) prolong the effects of both the preference shocks and the news shocks and (2) help capture the large volatility in the input expenditure shares. The main mechanism and the news transmission path will remain unchanged. How explicitly modeling inventory affects the results is not unambiguous. Intuitively, allowing for output inventory may dampen the price volatility hence reduces the effect of preference shocks. On the other hand, input inventory may amplify the effect of news, especially if the horizon of news is further than one period, because it allows the suppliers that are higher on the production chain to react to the news even earlier. While it can be interesting to explore, modeling inventory complicates the analysis significantly, and the amount of products that goes into the inventory is small on the aggregate level (for example, in 2007, the changes in private inventory accounts for less than 0.3% of total value added.) Now turn to the quantitative exercise. A natural next step is to consider natural experiments and conduct policy evaluation. For instance, it is useful to see how the industries react to an increase in military expenditure and/or to the news of that increase. More insights may be gained by augmenting the model to study international trades and studying the spill-over effects between countries.

Chapter 2

Endogenous Productivity Dispersion over the Business Cycles

2.1 Introduction

Cross-sectional productivity dispersion tends to rise in bad times. This is the case for productivity at the plant, firm, and industry level. Recently, this phenomenon has attracted growing attention from economists, with much new evidences from micro-level data sets.[1] However, the significantly negative correlation between uncertainty and aggregate economic conditions is often treated as a calibration discipline, and not much work has been done to explain it.

[1]Examples are Higson, Holly and Kattuman (2002), Higson, Holly, Kattuman and Platis (2004), Bloom (2009), Bloom, Floetotto and Jaimovich (2010), Bachmann and Bayer (2011), Arellano, Bai and Kehoe (2009), Bachmann, Elstner and Sims (2011), Chugh (2010), Kehrig (2011), to name a few.

In this paper, I provide a possible mechanism through which the worsened aggregate economic conditions lead to an increase in the measured dispersion in firm/plant-level productivity.[2] The model employed is close to the standard industry dynamics model with firm entry and exit built in the seminal work of Hopenhayn (1992), with aggregate fluctuations in "technology shocks" as the driving force of business cycles. Meanwhile, it differs from the standard model in that in each period, after observing the aggregate "technology shock realization," a staying firm has the option to adopt a risky project, in addition to a standard safe project whose productivity realization is determined by the aggregate state. Given the same capital input, the output and productivity associated with the risky project is a mean-preserving spread of the safe project's output and productivity. Although firms are risk neutral and the risky project does not give a higher flow payoff, there is a positive fraction of firms that strictly prefer to take the risk. This is because the option of exit provides a lower bound for a firm's continuation value as a function of working capital and creates a local convexity in it. Therefore, firms in this region have the incentive to randomize over their future values by choosing the risky project, and when the uncertain productivity is realized, dispersion arises. This setup resembles Vereshchagina and Hopenhayn's (2009) model of occupational choice. In bad times, this risky region gets larger and the fraction of risky firms rises. Consequently, the average or aggregate riskiness in firms' production increases, and so does the realized productivity dispersion. Despite the fact that the model is fairly standard with one little twist, it is capable of generating productivity dispersion negatively correlated with the aggregate state of the economy,

[2]This paper is not on firm theory. In what follows, the difference between a firm and a plant is not distinguished. The optimal number of plants/establishments a firm should have, although an interesting and important question, is not the focus.

with a correlation coefficient quantitatively in line with the data.

This model's mechanism is strongly motivated by empirical findings. It has features and implications that mirror the following micro-observations: (1) business cycle indicators lead the change in productivity dispersion; and (2) in recessions, more firms become risky and the exit rate is therefore countercyclical; (3) new firms are relatively small and small firms have a low survival rate; (4) small and/or young firms tend to bear more risk and/or show larger productivity dispersion.

The first two points involve the cyclical change. The increase in measured cross-sectional dispersion of plant- and/or firm-level productivity lags the worsened business cycle indicators, for example, the GDP growth rate, as shown in Bachmann, Elstner and Sims (2011) and Kehrig (2011) among others. A similar response is observed in the stock market. The last point relates to the key feature of the model. Although, unfortunately, I do not have direct observations from the data, there is indirect evidence that implies that there is a larger fraction of risky firms in recessions, consisting mainly of small firms. The exit rate rises in bad times. The findings on the relation between firm size and exit rate show that small firms and establishments drive the negative correlation between the exit rate and business cycles. This indicates that small firms are more sensitive to the cyclical change, as the model predicts. The increased exit rate in bad times is shown in papers such as Campbell (1998) and Jaimovich and Floetotto (2008) and is discussed in Section 2.2. Perhaps more direct evidence is found in the cyclical pattern of price dispersion recently documented in Bachmann and Moscarini (2011) and Berger and Vavra (2011). Cross-sectional dispersion in price changes is countercyclical, both within and across sectors. Meanwhile, the dispersion is positively correlated with the frequency of adjustments, which is also countercyclical. The higher adjustment frequency in bad times can be interpreted as a result of firms doing more frequent risky pricing experiments due to lower experimentation cost, as in Bachmann and Moscarini (2011).

The latter two points are closely related, as the exit hazard is a special form of firm-level risk. The relation between firm size and dynamics is well established and can be traced back to, for example, Dunne, Roberts, and Samuelson (1988). This is discussed further in Section 2.2. The findings on firm size and riskiness mainly come from two directions. First, it is well established in the entrepreneurship literature that entrepreneurs, especially poorer ones, bear a substantial amount of risk and tend to hold largely undiversified assets by investing heavily in their own firms, despite little or no premium in doing so. The risk here is interpreted as either the dispersion in small firm owners' personal income or the dispersion in return to private equity. At the same time, privately owned businesses are, on average, smaller in scale, measured in terms of either capital stock, number of employees, or output.[3] The second stream of empirical findings, more relevant to my work, regards the productivity and firm size differential. Gertler and Gilchrist (1994), using the Quarterly Financial Report for Manufacturing Corporations, find that smaller firms exhibit a higher standard deviation in sales growth rates than larger ones do. Dhawan (2001) looks at publicly traded firms in Compustat and finds that small firms have a higher failure rate and a larger standard deviation in profit rate, while, conditional on surviving, small firms show a higher average profit rate. The superior profitability in small firms is reducd if profits are adjusted according to

[3]Examples of work in this direction are Hamilton (2000), Moskowitz and Vissing-Jorgensen (2002), and Herranz, Krasa and Villamil (2009). See Quadrini (2009) for a detailed review.

the failure rates. Here, Dhawan defines the profit rate as operating income per unit of capital, and he defines the firm-level riskiness or volatility as the variance in the random realizations of production. Using his definitions, my model generates the same pattern of profit rate and riskiness differential by size. There is also evidence from outside the U.S. For example, using a firm-level German data set, USTAN, that covers the majority of German industries, Bachmann and Bayer (2011) find decreasing productivity risk in firm size, where the risk is measured as the average cross-sectional standard deviation in log-differences in firm-level Solow residuals.

The goal of this paper is to complement existing theories on what causes the negative correlation between business cycles and cross-sectional productivity dispersion. It is true that, if measured uncertainty and aggregate economic conditions are correlated, the cause can be from either direction. The real option literature that aims to explain such countercyclicality suggests the opposite direction for a causal relationship, from increased uncertainty to decline in aggregate economic activity. An influential paper in this area is Bloom (2009), which was later generalized by Bloom et al. (2010). Bloom shows that increased uncertainty, through the channel of adjustment costs to capital and labor, leads to a larger option value of waiting and a pause in investment and employment. A sizable drop in aggregate economic activity occurs because of this "wait-and-see" effect. The time-varying uncertainty is twofold in his model: (1) the time series standard deviation of productivity can be either high or low, evolving as a Markov process, and (2) the one-step-ahead conditional variance of this Markov process depends on current realization. However, Bachmann and Bayer (2011) and Bachmann, Elstner and Sims (2011) show that there is little evidence of sizable "wait-and-see" effects in the data. In addition, the process of entry and exit is neglected. Arellano, Bai and Kehoe (2009) do consider the entry and exit dynamics that interact with financial constraints, but, again, the causal direction is from a time series uncertainty shock to a sizable response in aggregate variables.

It is important to note that the importance of the uncertainty shock is not denied in this paper, and the inverted causality may still be true. But there is an issue regarding measuring uncertainty, which relates to the lead-lag relationship between uncertainty and cycles. Time series variances of major business condition indicators are often interpreted as uncertainty. In addition, a parallel family of uncertainty measures concerns the realized cross-sectional dispersion in micro-level performance, which includes, among other things, the cross-sectional variance in measured firm-level total factor productivities, levels or growth rates, and sales growth rates. However, realized cross-sectional dispersion is only a proxy for uncertainty. Besides, increased micro-level cross-sectional dispersion tends to lag recessions. This suggests a possible causality from an aggregate economic state to measured uncertainty, in particular, cross-sectional dispersion in productivities. This paper looks at this interesting issue from an angle different from the one adopted by the aforementioned literature.

The other paper that entertains the same causal direction as mine is Bachmann and Moscarini (2011). They build a model in which firms need to run costly experimentation and hence learn about their own market powers. As a result of lower experimentation costs, the dispersion of productivity measured in sales is larger during recessions due to more experiments being conducted. My model shares a similar feature with theirs, in that the option of exiting promotes the risky performance of firms, while the rest of the mechanism is very different. At the same time, my model differs from theirs by predicting that smaller firms are the major contributors to productivity dispersion and entry/exit dynamics.

The rest of the paper is organized as follows. Section 2.2 describes the stylized facts on the cyclical dispersion of productivity, firm size distribution, and dynamics. Section 2.3 contains a simple three-period model that illustrates the mechanism and shows preliminary results. Section 2.4 takes the simple model to an infinite horizon. Section 2.5 concludes.

2.2 Empirical Facts

Cyclical Productivity Dispersion. Eisfeldt and Rampini (2006) use data from Compustat and find countercyclical movement of dispersion in Tobin's q. At the same time, they show a similar pattern for dispersion of total factor productivity growth rates at the four-digit SIC level, with a correlation of -0.465. Bloom (2009) shows that U.S. stock market volatility, as measured by the VXO index, is positively correlated with the cross-sectional standard deviations of firm profit growth, firm stock return, and industrial total factor productivity (TFP) growth at the four-digit SIC level, but its correlation with industrial production is significantly negative. Moreover, Bloom, Floetotto and Jaimovich (2010) take an even closer look at this issue and examine the Census of Manufactures. They find that various measures of uncertainty are significantly countercyclical at all establishment, firm, industry, and aggregate levels. Bachmann and Bayer (2011) use a long panel of German firm-level micro-data that covers all single-digit industries to show that the correlation between dispersion in growth rates of firm-level TFP, sale, and value added and economic performance is significantly negative. This pattern is maintained in subsamples divided by sector and by size. Although taken from a different economy, their USTAN data set has the clear advantage in coverage. Moreover, by looking at different size quantiles, they document that average time series productivity dispersion in smaller firms tends to be larger than in bigger firms. Chugh (2010) explores the profitability series constructed by Cooper and Haltiwanger (2006) from the Longitudinal Research Database and calculates the cyclical correlation between average productivity and the dispersion of profitability to be -0.97. However, the sample is of relatively short length, covering only 1977-1988, a period that exhibits an unusually large degree of opposite movement. Kehrig (2011) focuses more on the dispersion of productivity levels rather than profit rates. He looks at the establishment-level data of the U.S. manufacturing sector that consists of the Annual Survey of Manufactures, the Census of Manufactures, the Plant Capacity Utilization Survey, and the Longitudinal Business Database. Although the manufacturing sector as a whole shows a countercyclical dispersion in establishment-level TFP, the durable goods industries show stronger cyclicality and it is the firms in the bottom quantile of the productivity distribution that drive the dispersion dynamics.

In this paper, I study how the aggregate economic state affects the dispersion in micro-level productivity. To link my model to data, ideally, the aggregate state is the average productivity measured as the cross-sectional average of plant-level TFP, and the dispersion is then the variance or inter-quantile range of plant-level TFP. Lacking the plant-level data, I use industry data at the four-digit SIC level to approximate the desired measures. The paper is silent on the validity of this

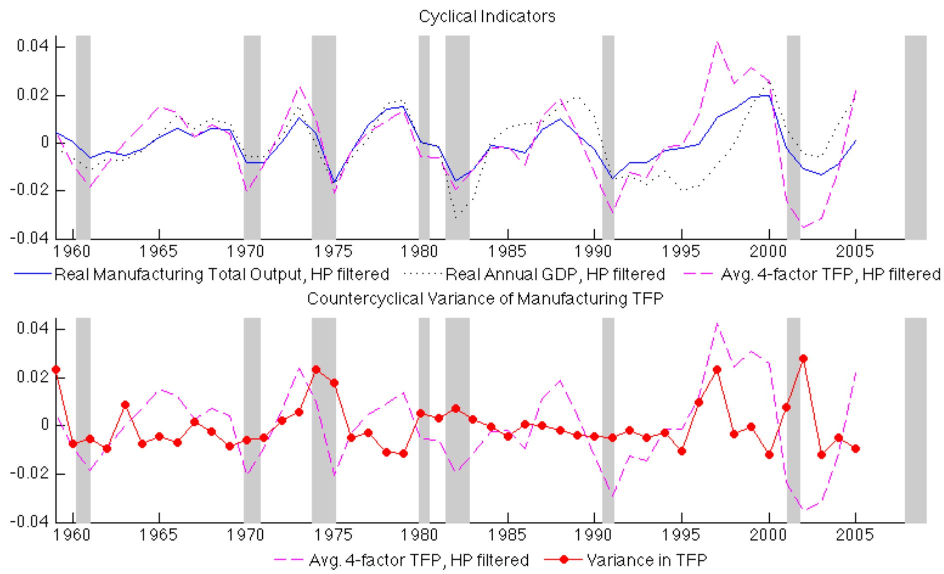

Figure 2.1: Cyclical Indicators and Variances in TFP. The upper panel plots different cyclical indicators, Real GDP (dotted line), Real total manufacturing output (solid line), Average TFP across industries at the 4-digit SIC level (dashed line). The lower panel shows the cyclical behavior of TFP dispersion measured as the variance (solid line with dots), together with Average TFP (dashed line). All series are HP-filtered. The shaded bars indicate official NBER recessions. Real GDP data are from FRED; TFP series are from MIPD, as is Manufacturing output measured as Real Total Shipments.

approximation, but Bloom et al. (2010) show that the countercyclical patterns of productivity dispersion are similar at the plant, firm, and industry levels.

The upper panel of Figure 2.1 shows the co-movement of different business cycle indicators. In particular, I claim that the average TFP is a valid aggregate state indicator for the manufacturing sector. The correlation coefficient between average TFP (HP filtered) and sectoral output (HP filtered) is 0.86 with a p-value of scale 10^{-9}. The average TFP corresponds to the cyclical indicator used throughout the model, and the fluctuation in TFP represents a technology or productivity shock, which drives the dynamics of the model economy. Following Eisfeldt and Rampini

(2006) and Bloom (2009), I use dispersion in the cross-sectional distribution of the TFP growth rate at the four-digit SIC level to approximate that at the individual level, without arguing the validity of the approximation. Note that the desired distribution is that of the *levels* of TFP instead of growth rates. The result is the lower panel of Figure 2.1, which illustrates the countercyclical movement of the variance in TFP.[4] The precise correlation coefficients for the U.S. manufacturing sector are documented in detail in both Bloom, Floetotto and Jaimovich (2010) and Kehrig (2011) and are summarized in Table 2.1 together with my own calculations.

Due to the limitations of the data, I use dispersion measures for the TFP growth rate instead of the TFP level. The corresponding cyclical indicators are then the GDP growth rate, the sectoral output growth rate, and the average TFP growth rate. To be comparable to other works, I include only the GDP growth rate and GDP HP residuals in Table 1.

Firm Dynamics. One important cyclical feature of firm dynamics that motivates this paper is that the exit rate moves countercyclically. This phenomenon is well documented in Campbell (1998) who uses ASM data between the second quarter of 1972 and the last quarter of 1988. In addition, Jaimovich and Floetotto (2008) assemble a new annual data set from 1956 to 1996 at the firm level across a broader range of industries and find that despite the difference in numbers, the exit rates of all examined industries are countercyclical. To illustrate firm dynamics over time, I obtain annual data from 1977 to 2009 in Business Dynamics Statistics (BDS) at CES, a data set that recently became publicly available. To

[4]I obtain data from the same sources as the aforementioned two papers, yet with more recent data up to 2005. I get the same significantly negative correlations as in those two papers if I use only the same range of data as they do. However, if I include the updated data as shown in Figure 2.1, I find a negative correlation that is not significant and is much smaller in absolute scale, which is less than 0.11.

Table 2.1: Correlations Between Dispersion and Cyclical Indicators

	U.S. Manufacturing	ΔGDP	GDP	Avg. ΔTFP
Kehrig (2011): Establishment level TFP Std. Dev.				
(1)	Durables	-0.420	-0.528	–
(2)	Non Durables	-0.172	–	–
Bloom et. al. (2010): Establishment (E) and firm (F) level.				
(3)	ΔOutput, IQR, E	-0.364	–	–
(4)	ΔTFP, Std. Dev., E	-0.273	–	–
(5)	ΔSales, IQR, F	-0.265	–	–
(6)	Stock Returns, IQR, F	-0.339	–	–
Calculated from NBER-CES MIPD: Industry level ΔTFP				
(7)	IQR	-0.502	-0.298	-0.184
		(0.000)	(0.021)	(0.108)
(8)	Std. Dev.	-0.262	-0.241	-0.129
		(0.038)	(0.051)	(0.194)
(9)	Var.	-0.249	-0.245	-0.123
		(0.046)	(0.048)	(0.206)

The first column of results shows the correlation coefficients (p-value) for real GDP growth rate, the second for residuals of HP-filtered real GDP, and the last for the weighted average TFP growth rate in the manufacturing sector. Rows (1) and (2) are taken from Tables 3 and 4 in Kehrig (2011), in which the micro-level data sources are mainly ASM/CM/LBD, continuously covering the period of 1972-2005 at an annual frequency. Rows (3) to (6) are from Table 1 in Bloom, Floetotto and Jaimovich (2010). Establishment-level data are also from ASM/CM/LBD, 1972-2006, while the firm-level infomation is from Compustat at quarterly frequency, 1967:II-2008:III for sales growth and 1969:I-2010:III for stock returns. Rows (7) to (9) are TFP dispersions across industries at the four-digit SIC level and the NBER-CES Manufacturing Industry Productivity Database is the source, covering 1959-2005 at an annual frequency. Except for the inter-quantile range (IQR), all other moments of industrial TFP growth are weighted by the real value of total shipments. Numbers in parentheses are one-sided p-values under the null of non-negative correlation.

be consistent with micro-level evidence on countercyclical dispersion, I look only at establishments in the manufacturing sector.[5]

Table 2.2 summarizes the establishment entry and exit rates by firm size.[6] A firm is classified as small if it has fewer than 50 registered employees. This is again not ideal, but subject to data availability. The preferred size classification is by capital stock. A more detailed illustration of entry and exit rates by year and by establishment size can be found in the Appendix.

Comparing establishment dynamics in small firms to those of large ones, they are of a much larger scale, more volatile, and more cyclical. Therefore, in the quantitative model, I focus only on the dynamics in small firms and treat the entry and exit of large firms mainly as exogenous, and they happen only with small probability.

The model I build in the following sections tries to explain the negative correlation between average productivity and cross-sectional productivity dispersion. The main mechanism emphasizes the different behavior between small and large firms,

[5] A noteworthy issue here is how to define an entering establishment and an exiting one. According to the official overview of the BDS data set, "An establishment opening or entrant is an establishment with positive employment in the current year and zero employment in the prior year. An establishment closing or exit[ing] is an establishment with zero employment in the current year and positive employment in the prior year. The vast majority of establishment openings are true greenfield entrants. Similarly, the vast majority of establishment closings are true establishment exits (i.e., operations ceased at this physical location). However, there are a small number of establishments that temporarily shutdown (i.e., have a year with zero employment) and these are counted in the establishment openings and closings." Therefore, an inevitable caveat is that although of relatively small number, an "idling" establishment can show up in the data first as an exiting one, and then as an entrant, for potentially many times. However, one clear advantage especially over firm-level data is that mergers and acquisitions are not reasons for disappearing units. Therefore, I can safely assume that exiting establishments suffer from low realizations of productivity.

[6] The entry and exit rates are indeed calculated using the numbers of newborn establishments, closed establishments, and existing establishments. However, the size is classified using the number of employees in a firm, instead of an establishment. One can only argue that large firms tend to own large establishments, and therefore large establishments exhibit similar dynamics to the ones owned by large firms. Otherwise, it is not clear whether this is a valid approximation.

Table 2.2: Entry and Exit Rates in Manufacturing Sector

U.S. Manufacturing 1977-2009		Total	Large	Small
(1)	Avg. Entry Rate (%)	9.36	5.18	31.18
(2)	Avg. Exit Rate (%)	9.28	6.00	30.06
(3)	Std. Dev. ($Entry^{HP}$) (%)	0.52	0.64	1.85
(4)	Std. Dev. ($Exit^{HP}$) (%)	0.67	0.90	1.56
(5)	Corr($Entry^{HP}$, $(\text{Avg. TFP})^{HP}$)	0.20	0.19	0.21
		(0.29)	(0.33)	(0.29)
(6)	Corr($Exit^{HP}$, $(\text{Avg. TFP})^{HP}$)	-0.26	-0.17	-0.23
		(0.17)	(0.37)	(0.24)
(5')	Corr(ΔEntry, Avg. ΔTFP)	0.22	0.13	0.31
		(0.26)	(0.51)	(0.11)
(6')	Corr(ΔExit, Avg. ΔTFP)	-0.10	0.06	-0.06
		(0.62)	(0.76)	(0.73)

The data source is still BDS. The binary grouping rule in size can be found in the caption for Figure 2.2. In Rows (1) and (2), the numbers are simple time series averages. Rows (3) and (4) are time series standard deviations for HP residuals. Rows (5) to (6) are correlations for HP residuals, and Rows (7) and (8) are for changes. Numbers in parenthesis are p-values. I choose to compute correlation coefficients in this way instead of using original entry/exit rates because there is a declining trend in both series. This is an interesting observation for its own sake, but this paper is silent on it.

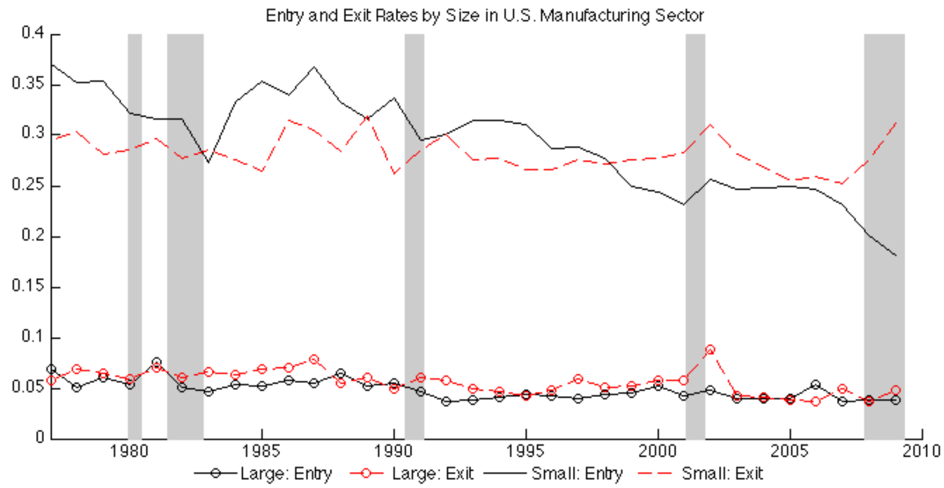

Figure 2.2: Cyclical Behavior of Entry and Exit in Manufacturing Sector by Size. A small firm is classified as one with fewer than 50 registered employees, and a large one with at least 50. This figure shows the original series of entry (solid lines) and exit (dashed lines) rates by size. The two thinner lines at the bottom are for large firms, and the two thicker ones are for small firms. Data on entry and exit rates are from BDS of CES.

which leads to observed differences in their entry and exit dynamics.

2.3 A Simple Model

To highlight the mechanism, I start from a simplified and tractable three-period version of the full model. I remove some features of the working model that are not as crucial and focus only on the incumbents' problem. The main idea is that the option to exit promotes risk taking of small firms by creating a local non-concavity in a firm's continuation value function, which in turn generates a non-degenerate dispersion in productivity. Moreover, as is shown in the comparative statics analysis, such dispersion becomes larger in bad times, due to a larger fraction of risk-taking firms. The same mechanism drives the infinite horizon model as well.

2.3.1 Setup

There are 3 periods, $t = 0, 1, 2$. There is a continuum of risk-neutral firm owners, each of whom owns a firm with different levels of initial resources $w_0 \in [0, \bar{w}]$. Assume that there is only one final good and each firm has only one plant that produces this good. The c.d.f. of owners' initial endowment of the single good is given as $G(w_0)$. At period 0, initial wealth w_0 can be divided into investment k_0 for future payoff and immediate consumption $w_0 - k_0$. If an owner decides to invest k_0, then she will get $w_1 = F(Z, k)$ as period 1 wealth, where

$$F(Z, k) = Zk^{\alpha}, 0 < \alpha < 1,$$

and Z represents the realized productivity of the project the firm owner chooses after the investment decision. A production project is associated with a project. Assume that production requires full attention of the firm's owner and uses the full capacity of the plant; hence, a firm cannot undertake multiple production projects simultaneously. An owner can choose one and only one out of two available projects: a safe one and a risky one, differing in the riskiness and realizations of productivity. For the safe project, $Z = A$ for sure, while for the risky one, with probability $p \in (0, 1)$, $Z = \bar{z} > A$, and with probability $1 - p$, $Z = \underline{z} = 0$. Both projects give the same expected value of Z, that is, $p\bar{z} + (1 - p) 0 = A$.[7] The risky project has a variance in productivity as a function of p and $\bar{z}$, $\sigma^2(p, \bar{z}) = p(1 - p)\bar{z}^2$. As a result of the linearity of $F(Z, k)$ in Z, the expected flow output of the risky project is the same as the safe one. Under this setup, A corresponds to the average

[7]For tractability, I assume only one type of risky technology and binary possible realization of it. In fact, a risky technology can be represented by a random variable Z with any distribution that is a mean-preserving spread of A.

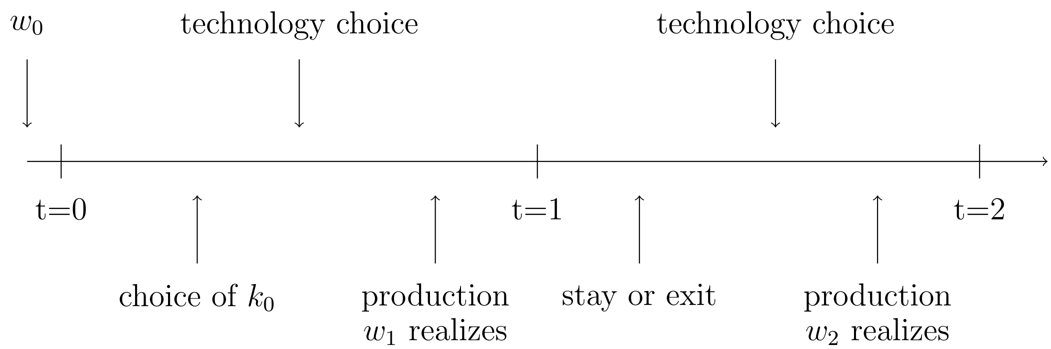

Figure 2.3: Timing of the Simple Model.

establishment-level productivity measured as TFP in the data and plays the role of economic condition indicator (or cyclical indicator in the full model); the riskiness of the risky project represents the risk at the establishment level, while its aggregated counterpart measures the dispersion in productivity.

2.3.2 Analysis

At period 1, after the uncertainty in Z is realized, the agent can decide whether to close her firm, exit the industry and get an outside option value V^0, or stay. Conditional on staying, she makes the investment choice k_1 and project choice again based on period 1 wealth w_1. In the last period, she simply consumes her final wealth w_2. The objective of an agent with initial wealth w_0 is to maximize her discounted consumption, with discount factor β:

$$V_0(w_0) = \max_{0 \le k_0 \le w_0} \left\{ (w_0 - k_0) + \beta \max \left\{ V_1(Ak_0^\alpha), (1-p) V_1(0) + pV_1(\bar{z}k_0^\alpha) \right\} \right\}$$

where $V_t(w_t)$ is the time t value for an agent with wealth w_t.

It is convenient to work backwards. At time $t = 2$,

$$V_2(w_2) = w_2.$$

At time $t = 1$, an agent with $k_1 > 0$ will be indifferent between operating a safe project and a risky one, simply because $Ak_1^\alpha = p\bar{z}k_1^\alpha$. Assume that all agents will perform safely in this case, which is consistent with their choice if they were risk averse. For simplicity, borrowing is not allowed, and the period 1 value for a staying firm will be:

$$V_1^1(w_1) = \max_{0 \le k_1 \le w_1} \{(w_1 - k_1) + \beta A k_1^\alpha\}.$$

Let k^* be the optimal capital choice for this firm, then

$$k^* = (\alpha\beta A)^{\frac{1}{1-\alpha}}.$$

The value of a firm with wealth level w_1 at the beginning of period 1 will be given by

$$V_1(w_1) = \max\left\{V^0, V_1^1(w_1)\right\}.$$

Let w_1^* be such that $V^0 = V_1^1(w_1^*)$. Note that there is a kink at w_1^* and $V_1(w_1)$ is convex in a neighborhood of w_1^*. This gives a firm with relatively low wealth level an incentive to take a risky project before it enters period 1. At $t = 0$, a firm makes the investment decision and chooses a project:

$$\begin{aligned} V_0(w_0) &= \max_{0 \le k_0 \le w_0} \{(w_0 - k_0) + \beta \max\{V_1(Ak_0^\alpha), (1-p)V_1(0) + pV_1(\bar{z}k_0^\alpha)\}\} \\ &= \max_{0 \le k_0 \le w_0} \left\{(w_0 - k_0) + \beta \max\left\{V^0, V_1^1(Ak_0^\alpha), pV_1^1(\bar{z}k_0^\alpha) + (1-p)V^0\right\}\right\}. \end{aligned}$$

To explicitly characterize a firm's project choice, it is useful to introduce the following condition on parameters.

Condition 1. $0 < V^0 < \alpha^{\frac{2\alpha^2}{1-\alpha^2}} \beta^{\frac{1+\alpha^2}{1-\alpha^2}} \bar{z}^{\frac{1}{1-\alpha}} p^{\frac{2\alpha^2}{1-\alpha^2}} \left(p^{1+\alpha} - p^2\right) / \left(1 - p\right).$

The risky and safe continuation values intersect at most once in the region where they are both greater than V^0. This condition ensures the existence of the intersection and makes the analysis tractable as shown in Proposition 3. The intuition is that given $(\bar{z}, p)$, the option value V^0 of exiting cannot be too high; otherwise, exit becomes very appealing, and so does the risky project. If the condition is violated, then all staying firms strictly prefer the risky project. In particular, if V^0 is given, this happens when A is low enough.

Proposition 3. *At $t = 0$, if Condition 1 holds, then the continuation value functions associated with risky and safe projects intersect only once, and $\exists k_0^I$ and k_0^{II} such that $0 < k_0^I < k_0^{II} < k^*$ and the decision rule of a firm's owner with initial wealth w_0 will be one of the following:*

1. *If $0 < w_0 \leq k_0^I$, she consumes all w_0 in period 0 and exits in period 1 for sure;*
2. *If $k_0^I < w_0 < k_0^{II}$, she invests all w_0 in a risky project in period 0, then with probability p, $w_1 = \bar{z} k_0^\alpha$, she in turn invests all w_1 in period 1; with probability $1 - p$, $w_1 = 0$, she exits in period 1;*
3. *If $k_0^I \leq w_0 \leq k_0^{II}$, she invests all w_0 in a safe project in period 0;*
4. *If $w_0 > k^*$, she invests k^* and consumes the rest in both periods.*

Figure 2.4 illustrates the agent's decision rule specified in Proposition 3. The interesting region, or the "risky region," is the interval $\left[k_0^I, k_0^{II}\right]$. The exiting option

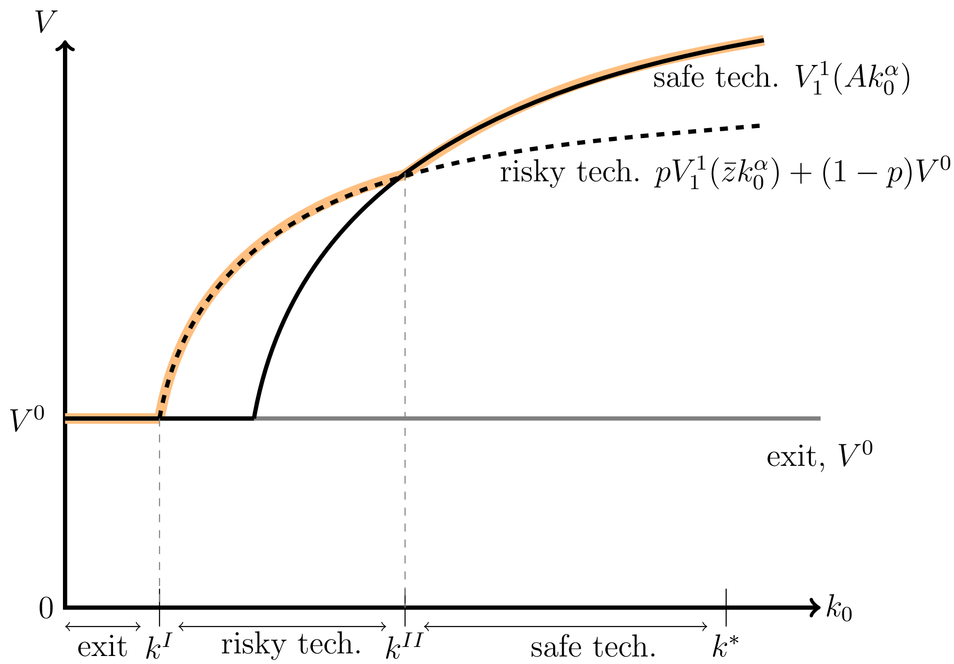

Figure 2.4: Continuation Values as Functions of Control Variable k_0. Illustration of Proposition 3. If $k_0 \leq k^I$, the wealth at time 1 is so small that the firm prefers to exit. If $k_0 \in (k^I, k^{II})$, the firm chooses risky asset. If $Z = \bar{z}$, $w_2 = \bar{z}k_0^\alpha$, if $Z = 0$, the firm exits and get V^0. If $k_0 \geq k^{II}$, the firm chooses the safe technology and obtains $V_1^1(Ak_0^\alpha)$ for certain. The highlighted portion is the maximized continuation value for each k_0, $\max\{V^0, V_1^1(Ak_0^\alpha), pV_1^1(\bar{z}k_0^\alpha) + (1-p)V^0\}$.

forms a lower bound in value function that is higher than in the case without exiting. This new lower bound alters the shape of the continuation value function, in particular, the continuation value function has a local convexity if safe project is chosen. This non-concavity region is roughly the same as the interval $\left[k_0^I, k_0^{II}\right]$, in which firms have a limited amount of capital stock. Firms that fall into this region have the incentive to smooth out such convexity by entering a lottery and randomizing over possible outcomes, which is exactly the role that risky project plays in this model. The fraction of risk-taking firms will then be determined given the initial distribution $G\left(w_0\right)$, and each of these firms bears the same risk in terms of the randomness of productivity.[8] As can be seen below, a change in A drives the changes in the risky region and the fraction of risk-taking firms and leads to a different productivity dispersion.

Suppose that with probability p the risky project is realized to have high productivity. The cross-sectional variance in realized productivity in period 0, denoted as $\Gamma\left(p, \bar{z}\right)$, is a function of p, the probability of good realization of the risky project, and $\bar{z}$, the good realization of productivity.

$$\begin{aligned} \Gamma\left(p, \bar{z}\right) &= E_{w_0,Z}\left(Z^2\right) - \left[E_{w_0,Z}\left(Z\right)\right]^2 \\ &= \sigma^2\left(p, \bar{z}\right)\Lambda\left(p, \bar{z}\right), \end{aligned}$$

where Z represents the productivity of the project a firm chooses, and $\Lambda\left(p, \bar{z}\right) := \frac{G\left(k_0^{II}\right) - G\left(k_0^I\right)}{1 - G\left(k_0^I\right)}$ in which k_0^I and k_0^{II} are functions of p and $\bar{z}$ as well. $\sigma^2\left(p, \bar{z}\right)$ is

[8]Once again, the same risk results from the assumption that only one way of randomization is permitted in the model for simplicity. To relax this restriction, one can assume that each firm can choose any distribution on productivity so long as the expectation remains A, which results in a risky region larger than $\left[k_0^I, k_0^{III}\right]$. However, while making the model much more complicated, this will not alter the result qualitatively, and neither will it provide more insight into the model.

simply the variance of the Bernoulli distributed productivity of the risky project, while $\Lambda(p, \bar{z})$ represents the measure of firms in the risky region. $\Gamma(p, \bar{z})$ is ex ante variance and coincides with realized dispersion in productivity, assuming a form of law of large numbers holds. At the same time, the aggregate or average output in period 0, $O(p, \bar{z})$, is:

$$\begin{aligned} O(p, \bar{z}) &= E_{w_0, Z}\left(F(Z, k_0)\right) \\ &= p\bar{z}\int_{k_0^I}^{k^*} w_0^{\alpha} dG(w_0|k_0 > 0) + p\bar{z}(k^*)^{\alpha} \frac{1 - G(k^*)}{1 - G(k_0^I)}. \end{aligned}$$

2.3.3 Comparative Statics

The nature of the simple model does not permit cyclical features. Therefore, I will instead analyze the comparative statics mimicking different times of business cycles. In particular, I use A, the average productivity, as the economic condition indicator, which corresponds to the average TFP in the data. In the model, a change in A can result from either a change in p, or in $\bar{z}$, or in both. Provided that the bad outcome of the risky project is normalized to be zero, $\bar{z}$ then determines the range, the variance of the Bernoulli productivity $\sigma^2(p, \bar{z})$, and the measure of the risky region $\Lambda(p, \bar{z})$. At the same time, $\sigma^2(p, \bar{z})$ and $\Lambda(p, \bar{z})$ are also nontrivial functions of p. When A, p, and/or $\bar{z}$ changes, the resulting change in the riskiness of a risky project, that is, variance $\sigma^2(p, \bar{z})$ or range $\bar{z}$, is called the "riskiness effect," as such change directly affects the riskiness of available project; and the change in the measure of firms in the risky region, $\Lambda(p, \bar{z})$, is the "mean effect," as the change in mean A determines the slope of continuation functions, which in turn affects the width of the risky region. The interesting one is the mean effect, which highlights the novel mechanism of the model; therefore, I consider a particular change in A, such that $\bar{z}$ is held unchanged and p is controlled for in a certain way to fully eliminate the riskiness effect, and I examine the resulting mean effect.

Proposition 4. *Let V^0 and $\bar{z}$ remain unchanged and assume Condition 1 always holds. Let $A \in \left\{A^H, A^L\right\} = \left\{p^H \bar{z}, p^L \bar{z}\right\}$, p^H and p^L be such that $p^H > p^L > 0$. Suppose the distribution of initial wealth $G(\cdot)$ is Pareto or uniform and the lower bound of its support is below k_0^I when the good outcome of the risky project is w.p. p^H. Then:*

1. $O\left(p^H, \bar{z}\right) > O\left(p^L, \bar{z}\right)$;
2. $\Lambda\left(p^H, \bar{z}\right) < \Lambda\left(p^L, \bar{z}\right)$.

   To control the riskiness effect, assume $p^H + p^L = 1$, then:

3. $\sigma^2\left(p^H, \bar{z}\right) = \sigma^2\left(p^L, \bar{z}\right) = \bar{z}^2 p^H p^L$;
4. $\Gamma\left(p^H, \bar{z}\right) < \Gamma\left(p^L, \bar{z}\right)$.

According to this proposition, given $\bar{z}$ fixed, A (or p) summarizes the aggregate state; higher A then means good times. When the aggregate state is good, the total output is high, and this is always the case whether the riskiness effect is controlled for or not. Meanwhile, the risky region is smaller in good times, which in turn leads to a smaller fraction of risk-taking firms, regardless of the riskiness effect. The assumption on Pareto or uniform distribution is not very restrictive. In fact, it can be any distribution that results in the same pattern of change in the fraction of risky firms. I choose Pareto distribution to mimic the actually observed size distribution of firms, which is only a sufficient but not necessary condition for the desired change in risky fraction. When the riskiness effect is controlled for, the riskiness of a risky project remains unchanged; therefore it is the change in the fraction of risk-taking firms that drives the change in resulting productivity dispersion, or the average riskiness that firms choose to take, measured as the variance in productivity.

If $\bar{z}$ is not fixed or p is not controlled for in such a way, then it is impossible to disentangle the mean effect from the riskiness effect, and these two effects jointly determine the resulting change in the cross-sectional dispersion in productivity. In fact, in the calibrated quantitative model, it turns out that the riskiness effect is too small to a generate significant difference in simulated results.

Figure 2.5 illustrates what happens to the model if A decreases, as described in Proposition 4. When A is low, the exiting threshold increases and more firms exit. At the same time, low A also leads to a larger risky region and a greater fraction of risk-taking firms; so now there are more firms that strictly prefer the risky project. As a result, if the change in A is controlled for as specified before, the average risk that firms choose to take is also larger and so is the realized productivity dispersion. To summarize, the key step for the model to generate a countercyclical productivity dispersion is the change in the risky region as the aggregate state changes. And it is mainly an enlarged fraction of risk-taking firms that causes a larger productivity dispersion in bad times. This mechanism remains in the quantitative model with infinite horizon. In fact, if the aggregate state follows a Markov process with only two possible outcomes of A^H and A^L controlled for in a similar way, then without introducing other features, the negative correlation between the aggregate state

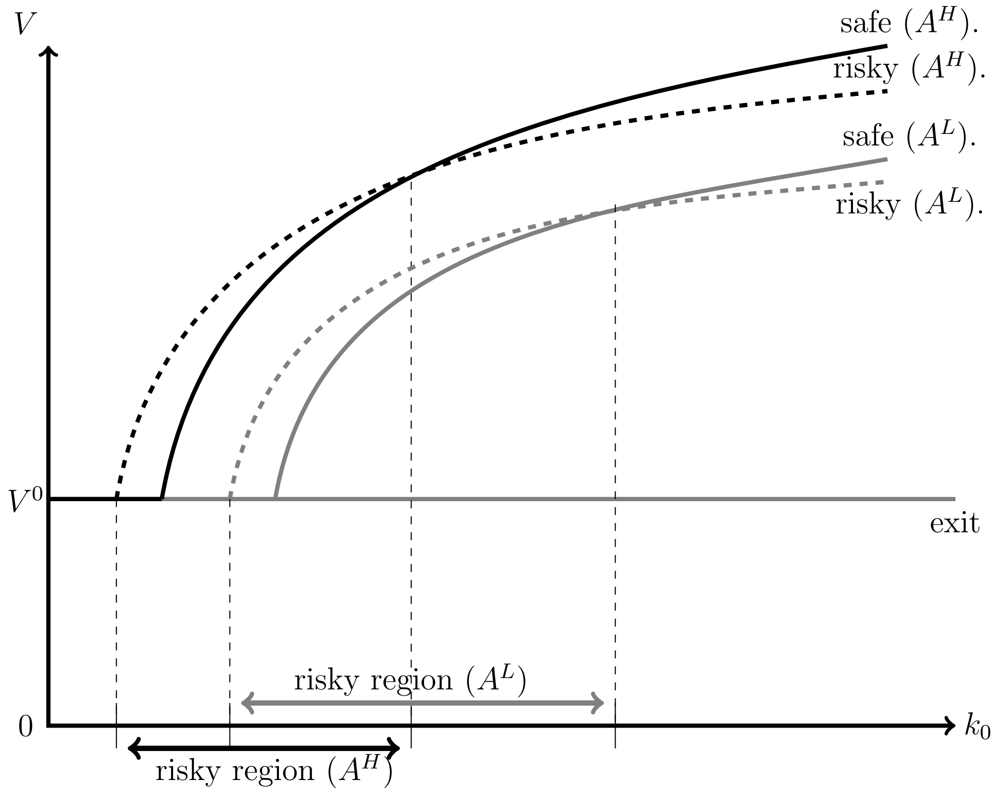

Figure 2.5: Comparative Statics. The continuation functions move downwards when the average productivity decreases from A^H to A^L.

and productivity dispersion is still almost perfect.

2.4 Quantitative Model

The simple three-period model illustrates the main mechanism in a tractable setting. However, it is only feasible to look at the comparative statics in an essentially static model with three stages. Therefore, a richer model with infinite horizon is built in this section to include more realistic business cycle features and to examine the quantitative performance of the mechanism.

2.4.1 Setup

Time is discrete, with infinite horizon. The firms that have survived at least one period are called incumbents. There is a constant mass $M > 0$ of potential entrant firms every period, each of which draws its initial capital k_0 from a distribution $G^0(k_0)$. $G^0(\cdot)$ determines the number and size distribution of newly born firms. Once it has entered, an entrant acts as an incumbent thereafter as long as this firm stays. The production function is the same as in the simple model, $F(Z,k) = Zk^{\alpha}$, with $0 < \alpha < 1$ and Z being the realized productivity depending on project choice. At the beginning of each period, all firms observe average productivity A. An incumbent firm owner makes the choice between staying and exiting; meanwhile, all firms also face an exogenous exiting probability $\eta > 0$. I allow additional exogenous exiting to generate the death of large firms, which always choose the safe project, as in the simple model. If an incumbent exits, the owner closes her firm and sells all capital stock. Once exiting, the firm cannot re-open for business again in the future. A staying firm then decides the amount of the next period's working capital k' and whether to adopt the safe project or the risky one. Again, assume the full attention of a firm's owner and complete utilization of plant capacity as a prerequisite of production. After production, capital depreciates at rate δ.

Under these settings, firms in this economy are heterogeneous in realized productivity, capital stock, and depreciation rate in each period; provided a realization of the aggregate state, project choice, investment, and depreciation jointly determine the incumbent's next period disposable resource.

The aggregate state for the model economy A evolves as a Markov chain with $A \in \mathbb{A} = \{A_1, ..., A_{N_A}\}$, and transition probability $\pi_{ij} = \Pr(A^j | A^i)$. In particular, this Markov chain is a discretized AR(1) process, such that $\ln A_t = \rho_A \ln A_{t-1} + \sigma_u u_t$, where $\rho_A \in (0, 1)$ is the serial correlation, and $u_t \sim N(0, 1)$ is white noise. Following conventional real business cycles models, I assume time-invariant volatility in A, in terms of constant σ_u. This implies that the driving force of this modelled economy is the traditional "technology shocks," that is, the change in the "first moment". This is different from Bloom (2009) and Bloom et al. (2010), who use time-varying higher moments as the pure source of aggregate fluctuation. Meanwhile, this is also distinct from, for example, Bachmann and Bayer (2011) and Chugh (2010), who allow time-varying higher moments in addition to the usual first moment movement to account for the countercyclical dispersion observed in the data. I do not allow σ_u to change over time based on the following considerations: (1) σ_u is time series volatility, which is not the same as the observed cross-sectional dispersion, (2) this model emphasizes a mechanism through which time-varying A generates realized productivity dispersion, and there is no need to introduce additional variation, and

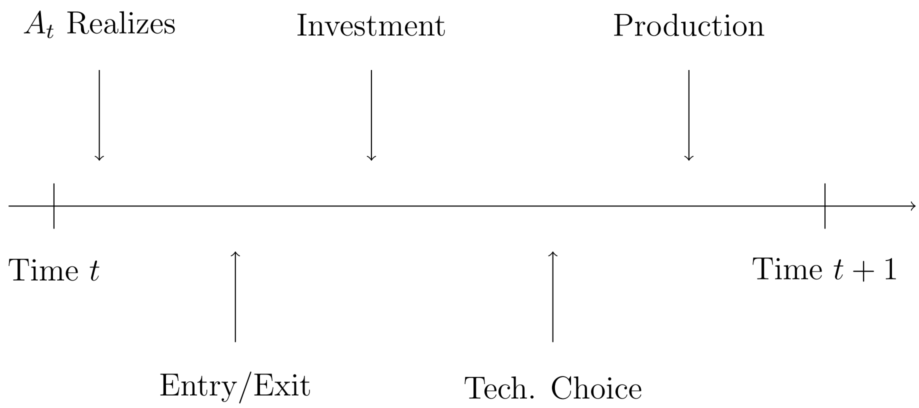

Figure 2.6: Timing of the Quantitative Model.

(3) fixed σ_u implies fixed unconditional mean of A.

Production is costly. In each period, a staying and active firm needs to pay a fixed operating cost, and if the firm needs to increase or decrease its capital stock, it pays a capital adjustment cost as well. Mainly following Cooper and Haltiwanger (2006) and Bloom (2009), I assume the capital adjustment cost consists of two parts: (1) a non-convex cost, and (2) a transaction cost. The non-convex cost represents the opportunity cost when a firm is under capital adjustment. Specifically, this firm forgoes a fraction c_k of its production if there is capital adjustment in a given period. The transaction cost represents the partial irreversibility. When a firm needs to increase capital, the price paid for every unit of new capital is normalized to be one, where the price is interpreted as how many units of output are needed to get one unit of capital. However, if a firm wants to reduce capital, the selling price for each unit of capital is $\theta < 1$.

Each time period has several stages, which resembles period 1 in the simple three-period model.

- Stage 1: Observation of state variables. Aggregate state A is realized. An incumbent firm observes A and enters this period with remaining capital after depreciation, $(1-\delta)\,k$, and together with last period's production $F(Z_{-1},k)$, where Z_{-1} is the realization of last period's productivity of this firm. A potential entrant draws k_0 and observes A.

- Stage 2: Entry and exit. An entrant with (k_0, A) enters if there is positive expected profit. An incumbent exits either voluntarily based on continuation values or exogenously with probability η.

- Stage 3: Investment and project decision. Both staying incumbents and newborn firms decide how much to invest and then choose between safe and risky projects. At the same time, the operating cost and capital adjustment cost are paid.

- Stage 4: Production. Production takes place in the form $F(Z,k')$, where k' is the new working capital, and Z is productivity. If a firm chooses the safe project, then productivity is deterministic, $Z=A$. Otherwise, with probability $p(A)$, the risky project turns out to be a success, $Z=\bar{z}$, and with probability $1-p(A)$, it fails, and $Z=0$.

2.4.2 Individual Decision

An Incumbent's Problem. At the beginning of each period, an incumbent firm is characterized by (Z_{-1},k,A), where $Z_{-1}\in\{A_{-1},0,\bar{z}\}$ is the realized productivity in the last period for a specific firm, which can be the safe productivity A_{-1}, the bad realization 0, or the good realization $\bar{z}$; k is the total amount of capital that was used in the previous period, and A represents the economic conditions of the current period.[9]

The first choice an incumbent firm owner makes is between continuing to operate and closing the firm and leaving.

$$V(Z_{-1}, k, A) = \max (1-\chi) V^1 (Z_{-1}, k, A) + \chi V^0 (Z_{-1}, k, A),$$

where $\chi \in \{\eta, 1\}$ is the exiting choice, and η is the exogenous exiting hazard. If a firm with (Z_{-1}, k, A) chooses to exit, the value is:

$$V^0 (Z_{-1}, k, A) = \theta (A) (Z_{-1} k^{\alpha} + (1-\delta) k);$$

where $\theta(A) < 1$ is the fraction of resources a firm owner can take away when exiting, which is actually a resale price and is potentially a function of A. If this firm chooses to stay, the owner must then decide on investment, i, and a project choice, safe or risky. The capital stock evolves as follows

$$k' = (1-\delta) k + i,$$

such that $k' \geq k_{\min} > 0$, where $k_{\min}$ is a very small positive number providing a lower bound of capital stock. The operating cost $C(i; Z_{-1}, k, A)$ of a firm consists

[9] To avoid computational complexity, I do not consider the price feedback effect in this model. Therefore, the distribution of firms is not a state variable in this model, because agents do not need to forecast future prices using information on distribution. Section 2.4.6.2 contains informal discussion on this matter.

of a fixed cost c_f and a capital adjustment cost:

$$C(i;Z_{-1},k,A)=c_f+c_kF(Z_{-1},k)\,1_{\{i\neq 0\}}+(1-\theta(A))(-i)\,1_{\{i<0\}}.$$

Apart from the fixed operating cost, there are two forms of capital adjustment costs: a non-convex adjustment cost and partial irreversibility. Actively adjusting capital stock and choosing $i\neq 0$ costs a firm c_k fraction of its revenue from the last period's production. In addition, if a firm reduces its scale, it can only sell its current capital possession at price $\theta(A)<1$. The fixed operating cost is to generate endogenous exiting behavior, and therefore, it creates a non-concave portion in the lower end of a firm's value function. The adjustment cost plays a double role: one is to capture the observed inaction in investment and slow down the change in firm size, and the other is to dampen firms' reaction to changes in aggregate states so that the correlation between productivity dispersion and the aggregate state is not too close to -1. Combining these pieces gives the flow profit of this firm $D(k';Z_{-1},k,A)$, and

$$P(i;Z_{-1},k,A)=F(Z_{-1},k)-i-C(i;Z_{-1},k,A)\geq 0.$$

I enforce non-negative profit as a constraint. The firm also has to choose between a safe and a risky project. A safe project produces $F(A,k')$ for sure; a risky project results in productivity at $\bar{z}$ with probability $p(A)$ and 0 with $1-p(A)$. If the safe one is chosen, the firm gets:

$$V^1_{safe}(i;k,A)=\mathbb{E}_{A',\delta'}\left[V(A,k',A')\,|A\right],$$

and likewise,

$$V^1_{risky}\left(i;k,A\right) = p\left(A\right)\mathbb{E}_{A'}\left[V\left(\bar{z},k',A'\right)|A\right] + \left(1-p\left(A\right)\right)\mathbb{E}_{A'}\left[V\left(0,k',A'\right)|A\right].$$

Therefore, conditional on staying, an incumbent firm solves the following maximization problem:

$$\begin{aligned} V^1\left(Z_{-1},k,A\right) &= \max_i\{P\left(i;Z_{-1},k,A\right) \\ &+ \beta\max\left\{V^1_{safe}\left(k';Z_{-1},k,A\right),V^1_{risky}\left(k';Z_{-1},k,A\right)\right\}\}. \end{aligned}$$

Denote the state variables of an incumbent as $\psi = \left(Z_{-1},k,A\right) \in \Psi$, with Ψ being the set of all possible states. The solution to an incumbent's question with state ψ is a list of policy functions $\{\chi\left(\psi\right),\tau\left(\psi\right),\iota\left(\psi\right)\}$ such that (1) $\chi\left(\psi\right)$ is the exiting choice, $\chi : \Psi \to \{\eta,1\}$; and conditional on surviving, (2) $\tau\left(\psi\right)$ is the project choice, $\tau : \{\psi \in \Psi : \chi\left(\psi\right) = \eta\} \to \{0,1\}$, where 0 represents the safe project and 1 the risky one, and (3) $\iota\left(\psi\right)$ is the investment level, $\iota : \{\psi \in \Psi : \chi\left(\psi\right) = \eta\} \to \mathbb{R}$.

A Potential Entrant's Problem. A potential entrant draws initial capital holding k_0 from a invariant Pareto distribution $G^0\left(k_0\right)$ with parameter ξ. The value of staying outside the market is

$$V^0_0\left(k_0,A\right) = \theta\left(A\right)k_0.$$

To start up a business, one must pay a setup cost c_e from initial capital, and thereafter acts as an incumbent with state (Z_{-1}, k, A) being

$$\psi_0 = (0, (k_0 - c_e) / (1 - \delta), A).$$

Hence, the payoff of opening a firm will be:

$$V_0^1(k_0, A) = V^1(0, (k_0 - c_e) / (1 - \delta), A).$$

A new firm will be born if

$$V_0^1(k_0, A) > V_0^0(k_0, A).$$

The solution to this problem is a binomial entry choice $\varepsilon : \Psi_0 \subset \Psi \to \{0, 1\}$, where Ψ_0 contains all possible ψ_0, and $\varepsilon(\psi_0) = 1$ if an entrant enters and 0 otherwise.

2.4.3 Aggregate Dynamics

Given the solutions to the individual problems described before, as a list of functions, $\{\chi(\cdot), \tau(\cdot), \iota(\cdot); \varepsilon(\cdot)\}$, it is straightforward to write down the transition dynamics for the distribution over $\psi = (Z_{-1}, k, A)$.

For an arbitrary $\psi \in \Psi$, either $\psi \in \Psi_0$ or ψ can only be the state of an incumbent. I denote $\phi(\psi)$ as the measure or density of point $\psi = (Z_{-1}, k, A)$ at Stage 1 of a typical period with aggregate state A, before entry and exit takes place. If $\chi(\psi) = 1$, then a firm with this state exits for sure, and no other transition can happen. If $\chi(\psi) = \eta$, then with probability η this firm exogenously exits, and with a complementary probability, it stays. Conditional on staying, if the firm chooses the safe project, $\tau(\psi) = 0$, then its individual state becomes $(A, (k + \iota(\psi)))$. On the other hand, if the firm chooses the risky project, $\tau(\psi) = 1$, then with probability $p(A)$ its individual state becomes $(\bar{z}, (k + \iota(\psi)))$, and with probability $(1 - p(A))$ it becomes $(0, (k + \iota(\psi)))$. Now turn to the newborns. Denote $g^0(\psi_0)$ as the entrant's measure or density at point ψ_0 determined by $G^0(\cdot)$. A newborn with ψ_0 enters if $\varepsilon(\psi_0) = 1$. After entering, this firm acts exactly the same as a surviving incumbent with $\psi = \psi_0$. Finally, the aggregate state becomes A' with probability $\Pr(A'|A)$, $A' \in \mathbb{A}$. Formally, suppose the aggregate state at Stage 1 of a period is $A' = A_j$, and that of the last period is $A = A_i$, meaning that the realized productivity Z is one of $\{A_i, \bar{z}, 0\}$. Every state not on the realization path has zero measure, or

$$\phi'(A, k', A') = 0 \text{ if } A \neq A_i \text{ or } A' \neq A_j,$$

where primed variables are ones realized in the same period as A'. The rest of the states can then be divided into three groups by realization of Z, all of which come from both incumbents and newborns. For $Z = A_i$,

$$\begin{aligned}\phi'(A_i, k', A_j) &= \int (1 - \chi(\psi))(1 - \tau(\psi))\, \mathbf{1}_{\{\psi : k' = (1-\delta)k + \iota(\psi)\}} \phi(d\psi) \\ &\quad + M \int \varepsilon(\psi_0)(1 - \tau(\psi_0))\, \mathbf{1}_{\{\psi_0 : k' = (1-\delta)k_0 + \iota(\psi_0)\}} g^0(d\psi_0),\end{aligned}$$

where variables with no prime are the ones observed one period back, with $\psi = (Z_{-1}, k, A_i)$ and $\psi_0 = (0, (k_0 - c_e)/(1 - \delta), A_i)$. For $Z = \bar{z}$ or 0,

$$\begin{aligned}\phi'(\{\bar{z}, 0\}, k', A_j) &= \int (1 - \chi(\psi))\, \tau(\psi)\, \mathbf{1}_{\{\psi : k' = (1-\delta)k + \iota(\psi)\}} \phi(d\psi) \\ &\quad + M \int \varepsilon(\psi_0)\, \tau(\psi_0)\, \mathbf{1}_{\{\psi_0 : k' = (1-\delta)k_0 + \iota(\psi_0)\}} g^0(d\psi_0).\end{aligned}$$

By independence, a fraction $p(A_i)$ has $Z = \bar{z}$, and the rest gets $Z = 0$, that is,

$$\begin{aligned}
\phi'(\bar{z}, k', A_j) &= p(A_i)\,\phi'(\{\bar{z}, 0\}, k', A_j),\\
\phi'(0, k', A_j) &= (1 - p(A_i))\,\phi'(\{\bar{z}, 0\}, k', A_j).
\end{aligned}$$

Given the distribution measure ϕ and ϕ', the cross-sectional variance in productivity can be written as

$$\begin{aligned}
\Gamma(A, \phi) &\propto \int \bar{z}^2 \phi'(\bar{z}, dk', dA') + \int A^2 \phi'(A, dk', dA')\\
&- \left[\int \bar{z}\phi'(\bar{z}, dk', dA') + \int A\phi'(A, dk', dA')\right]^2\\
&= \bar{z}^2 p(A)(1 - p(A)) \int \phi'(\{\bar{z}, 0\}, dk', dA') = \sigma^2(A)\,\Lambda(A, \phi).
\end{aligned}$$

The expression of the cross-sectional variance can be simplified in this way due to the linearity of productivity in production function.

2.4.4 Calibration

Before I describe the calibration procedure, it is worth noting that the mass of potential entrants M affects only the scale of the economy once other parameters are determined. Since the absolute scale is not of interest, the choice of M is irrelevant. For a quantitative exercise, the number of potential entrants is fixed at 50,000 each period. Furthermore, without aggregate fluctuation, starting from zero incumbents, the economy always converges to a stationary state in the sense that the exit rate and the entry rate are equal and the scale is neither expanding nor shrinking, as long as there is positive measure of entrants at the beginning. And this is the case with or without agents expecting the aggregate state to be varying over time. The reason is simple. Since there is no aggregate fluctuation, the measure of entrants (inflow) is fixed each period. The measure of exiting firms (outflow) is a fraction of the remaining ones (stock). The outflow gradually increases to the same level as the inflow, and it is at this point that the scale of stock stops changing. Consequently, the entry and exit rates are the same. Because of this stationarity feature, the parameters that need to be internally determined are selected such that the statistics generated by the model at its stationary state match their empirical targets.

The setup of the model is very close to that of the standard model; therefore some of the parameter values are directly taken from the literature. One period is chosen to be one year. The discount factor is set as $\beta = 0.938$ to match the long-run average for the U.S. firm-level discount rate, as in Bloom (2009). According to the same paper, capital depreciates at rate $\delta = 0.1$. The production function, $F(Z, k) = Zk^{\alpha}$, is the same as the profit function in Cooper and Haltiwanger (2006), so I follow their estimation and set α to be 0.592. Taken from the same work, the standard deviation of the aggregate process σ_A is 0.08, and the serial autocorrelation ρ_A is assumed to be 0.8, which is within the range of the autocorrelation of a common shock 0.76 and that of an idiosyncratic shock 0.885 estimated in that paper.

The good productivity realization is predetermined as $\bar{z} = 2$ so that the probability of getting $\bar{z}$ is always around a half. This is to minimize the riskiness effect by controlling for the uncertainty associated with the binary-outcome risky project. The exogenous exiting hazard η that affects all firms alike is set to be 2%, which is in line with the exiting rate of large plants found by, for example, Lee and Mukoyama (2008). On the entrant side, it has been mentioned that the choice of M is not important. The distribution of the initial endowment G^0 is Pareto such that, with slight abuse of notation, $G^0(k_0) = 1 - (k_{\min}/k_0)^{\xi}$ with $\xi > 0$. Clearly, ξ governs the shape of the initial endowment distribution and it in turn determines the model-generated firm size distribution. Ideally, this generated distribution will also have a shape close to Pareto; however, the assumption of one common productivity shock and no idiosyncratic shocks makes this task infeasible. This can be corrected by introducing heterogeneous productivity, yet this practice will not provide more economic insight into this model. Therefore, for the numerical results, I set $\xi = 1$.

The remaining parameters to be internally calibrated are capital resale price θ, capital adjustment cost as a fraction of profit c_k, fixed operating cost c_f, and entry cost c_e. The model suggests that I shall look at the statistics of firm dynamics and the investment rate distribution, and the remaining parameters (θ, c_k, c_f, c_e) are selected via simulated method of moments. The targets regarding firm dynamics are taken from Lee and Mukoyama (2008), and those on investment rate distribution are from Cooper and Haltiwanger (2006). I also compute from the model the average five-year transition rates between different size classes, and I compare the generated numbers to the actual rates found by Acemoglu, Akcigit, Bloom, and Kerr (2011) using census data. The parameters are calibrated without aggregate fluctuation, and the aggregate state sequence, $\{A_t\}$, is set to be constant at its mean, but the firms still expect the future states to be changing according to the transition probability of A, π_{ij}.

Calibrated parameter values are summarized in Table 2.3, [10] and simulated

[10] Several other sets of parameters are also used to check the robustness. The negative sign

Table 2.3: Parameter Values and Rationale

Parameters	Description	Notes
Aggregate Fluctuation		
$\bar{z} = 2$	Good realization.	Predetermined. Normalization.
$\rho_A = 0.8$	Autocorrelation.	Cooper and Haltiwanger (2006)
$\sigma_u = 0.048$	Innovation var. ($\sigma_A = 0.08$.)	Cooper and Haltiwanger (2006)
Production		
$\alpha = 0.592$	Production function.	Cooper and Haltiwanger (2006)
$\beta = 0.938$	Discount factor.	Bloom (2009)
$\delta = 0.1$	Capital depreciation rate.	Bloom (2009)
$\eta = 0.02$	Exog. exiting prob.	Lee and Mukoyama (2008)
$\theta = 0.84$	Capital resale price.	Internally determined.
$c_f = 1.62$	Fixed operating cost.	Internally determined.
$c_k = 0.165$	Capital adjustment cost.	Internally determined.
Entrants		
$c_e = 0.1$	Entry cost.	Internally determined.
$\xi = 1$	Shape of G^0.	Predetermined.

moments are compared with their empirical counterparts in Table 2.4. Cooper and Haltiwanger (2006) compute a thorough set of investment moments using a balanced panel from the LRD from 1972 to 1988. The model-generated moments are close to their target with expected exceptions. The standard deviation in investment rates is much lower than in the data, because when the aggregate fluctuations are shut down, there is no idiosyncratic uncertainty other than the amount of risk a firm chooses to take. With a constant aggregate state and no growth, the model-generated mean level of the investment rate, together with the fraction of

of the correlations between aggregate state and dispersion measures is robust, which is not surprising because the mechanism works under mild restrictions of parameter space. However, it is true that the fraction of risky firms is sensitive to the shape of the value function. In particular, when β is high, future profit flows are important, and the risky fraction declines and so does the exit rate. The realizations of δ are set to be $\{0.05, 0.1, 0.2, 0.5, 1\}$ with probabilities $\{0.69, 0.155, 0.1, 0.05, 0.005\}$, respectively.

large and positive investment rates, is below the target as well. The other set of targets concerns the entry and exit dynamics of firms, which are taken from Lee and Mukoyama (2008).[11] They use the ASM portion of the LRD from 1972 to 1997 to analyze the behavior of plants. At the same time, I look at the five-year transition rates between different size classes obtained by Acemoglu et al. (2011) using the CM portion. Firms are divided into two size classes, small and large, by median shipments, and the third class is "not-in-business." For example, the transition rate from the small class to the large class is computed as the fraction of originally small firms that became large ones in the next census. Since the census data are only available every five years, I let the model produce the same transition rates for every five periods. Due to different sources of data, I choose to hit a number within the range of empirically computed entry and exit rates. The model failed to reproduce the eight transition rates, although it managed to capture the fact that small firms have higher exiting rates than large ones. Without assuming idiosyncratic shocks, the model cannot generate a highly right-skewed size distribution with a relatively small median; therefore, the simulated exit rate is lower. At the same time, no further heterogeneity causes the large transition rates between large and small classes.

2.4.5 Quantitative Results

The mechanism explained in the illustrative three-period model remains at work in the quantitative model with infinite horizon. The option to exit forms a lower bound for an incumbent's continuation value function, and in a conventional model

[11] Lee and Mukoyama (2008) calculate the relative sizes of entering and exiting firms based on the number of employees.

Table 2.4: Moments Generated from Model and Targets

	Model	Data
Investment		
Mean of investment rate	0.097	0.122
Std. Dev. of investment rate	0.157	0.337
Fraction of inaction	0.059	0.081
Fraction w. positive investment	0.889	0.815
Fraction w. positive investment burst	0.064	0.186
Fraction w. negative investment burst	0.033	0.018
Data Source: Cooper and Haltiwanger (2006)		
Entry and Exit		
Mean entry rate	0.070	0.062
Mean exit rate	0.070	0.055
Relative size, entering	0.75	0.60
Relative size, exiting	0.58	0.49
Data Source: Lee and Mukoyama (2008)		

without the additional choice of risky project, this lower bound in turn creates a non-concave portion on the continuation value at the lower end with low capital levels. When the choice of risky project is allowed as in this model, firms with capital levels in this portion have an incentive to smooth out the non-concavity by taking the risk. Of course, anticipating the future option of the risky project, the continuation value function associated with the safe one becomes less convex compared to the conventional case.

The business cycle features can now be introduced in a more realistic fashion than comparative statics. Without recalibrating, I add the aggregate fluctuation by simulating a sequence of realizations of productivity level A, and let the model evolve accordingly. As the aggregate state changes, the reaction of firms is still very similar to the comparative statics in the simple model. If A drops, the slopes of both risky and safe continuation value functions decrease, which forms a larger portion where the risky project is strictly better. Consequently, a larger fraction of firms opt to take the risk, which results in a larger cross-sectional standard deviation in productivity. The opposite happens when A increases. Nonetheless, given the frictions and the law of motion of the aggregate state, the magnitude of the changes in the fraction of risk taking firms and in the resulting standard deviation in productivity is history dependent.

The main goal of this numerical exercise is to show that changes in the level of A_t alone can generate countercyclical firm-level productivity dispersion as a result of a firm's risk-taking behavior, without introducing any time-varying volatility in the driving force, A_t. The fluctuation in productivity A follows the Markov process specified in Table 3, and not surprisingly, it is positively correlated with total output with correlation coefficient 0.4030 (p-value = 0.0000). Therefore, the cross-sectionally averaged productivity can serve as an alternative cyclical indicator. The measures for productivity dispersion are chosen to be (1) the standard deviation of cross-sectional distribution of realized Z, productivity, (2) the fraction of firms that prefer the risky project, and (3) the 95% to 5% inter-percentile range of realized Z, which is the value of Z at the 95th percentile minus the value of Z at the 5th percentile.

Table 2.5 shows that the correlation coefficients between productivity dispersion and cyclical indicators are significantly negative, and the absolute values are in line with the data counterparts. In fact, the correlation between productivity dispersion and total output is even larger in scale. Moreover, the cyclicality of productivity dispersion measured is on a scale comparable to that of the fraction of

Table 2.5: Generated Cyclicality

Cyclicality: Correlations (p-value) with Cyclical Indicators

Variables of Interests		Cyclical Indicators: Avg. Productivity, A	Cyclical Indicators: Total Output, O
Productivity Dispersion	$\sigma(Z)$	-0.4450 (0.0000)	-0.6969 (0.0000)
Frac. of Risky Firms	Λ	-0.4544 (0.0000)	-0.6063 (0.0000)
IPR (95%-5%)	IPR_5^{95}	-0.2089 (0.0000)	-0.6860 (0.0000)
Entry Rate	r^{EN}	0.0314 (0.4830)	-0.7679 (0.0000)
Exit Rate	r^{EX}	-0.4774 (0.0000)	-0.5649 (0.0000)

firms that choose the risky project, and the movements show patterns very similar to those seen in Figure 2.8. This illustrates the mechanism that it is the change in the fraction of risk-taking firms that drives the cyclical movement of productivity dispersion. In bad times, more firms are willing to take the risk and randomize their future values. Consequently, the resulting dispersion, measured as the standard deviation of cross-sectional productivity distribution, is larger and so is the inter-percentile range.[12] The assumed binomial outcome of a risky project has the potential to impact the behavior of the dispersion; however, such impact is controlled for at a much smaller scale by the choice of μ_A and $\bar{z}$ and does not alter the main pattern. A somewhat unusual result is the significantly negative correlation between total output and entry rates. This is a result of modeling technique. The entry decision of potential entrants depends largely on the discounted and expected

[12]Due to the model assumption, cross-sectional IPR in productivity can only be either $\bar{z}$, $\bar{z}-A_t$, or A_t, and does not have very interesting dynamics, although it is still countercyclical. This can be overcome by allowing a richer set of productivity lotteries and keeping the expected productivity to be A. For example, in addition to $(p(A), \bar{z})$, firms can also choose any $(p, \bar{z}_A)$ pair with binary outcomes such that $p\bar{z}_A = A$. Intuitively, the IPR measure in this case will again be negatively correlated with A_t because smaller firms have the incentive to take even more risk in bad times than in the original case. Therefore, the range of realized productivities is wider, and potentially the IPR is larger and has more possible values.

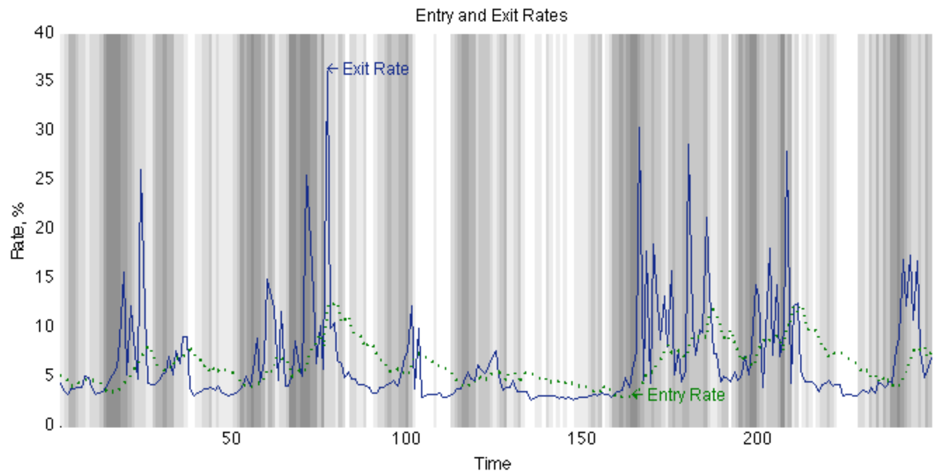

Figure 2.7: Simulated Sequences of Entry and Exit Rates. The solid line represents exit rates, and the dashed line records entry rates. Grey bars indicate the value of A as in the previous figure.

future payoff, so the impact of the current aggregate state is minimal. At the same time, entry rates increase when the number of existing firms is smaller. However, the total output is not only a function of the current state A, but it also positively depends on the number of existing firms. These two forces drive the entry rate series to move in the opposite direction to total output.

Figure 2.7 plots the truncated series of entry and exit rates from the model simulation. The sequence of exit rates remains mostly in a reasonable scale between 3% and 12%. On the contrary, there are quite a few episodes in which exit rates are really high. Extraordinarily high exit rates happen after a succession of bad realizations of the aggregate state A, when the number of remaining firms is small. This is not surprising under the model assumptions that (1) all firms share the same serially correlated A with no idiosyncratic shocks, and (2) given each realization of A, there is only one alternative risky project allowed. Figure 2.8 shows the truncated sequences of the countercyclical cross-sectional standard deviation in

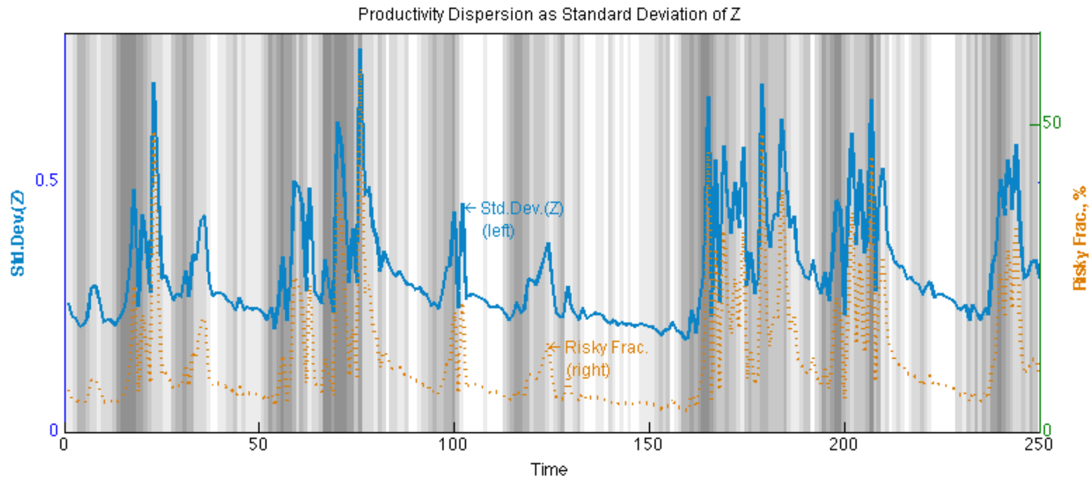

Figure 2.8: Simulated Sequences. The figure shows simulated sequences of (1) cross-sectional productivity dispersion measured as the standard deviation of realized productivity Z (solid line, left axis), and (2) the fraction of firms that choose the risky technology (dotted line, right axis, in %). The grey bars indicate the economic conditions as a value of A. In particular, darker bars represent lower values of A.

productivity and the fraction of risk-taking firms in each period. The realized standard deviation in productivity mostly ranges from 0.25 to 0.65, and the fraction of firms choosing the risky project is mostly between 10% and 55%. The peaks of productivity dispersion and the risky fraction are associated with excessive exit rates, as the mechanism suggests.

Figure 2.9 shows how the productivity dispersion and fraction of risk-taking firms will react to a drop in A from its mean level. Originally, the model is simulated in the same way as it is for calibration: the aggregate fluctuation is shut down by fixing A at its mean level μ_A, while the firms behave under the belief that A evolves according to π_{ij}. Then, the value of A suddenly and permanently switches to one standard deviation lower, $\mu_A - \sigma_A$, and the firms' belief remains unchanged. The risky fraction and productivity dispersion increase immediately upon impulse, then oscillate with an ascending trend, and eventually settle at a higher level. The two

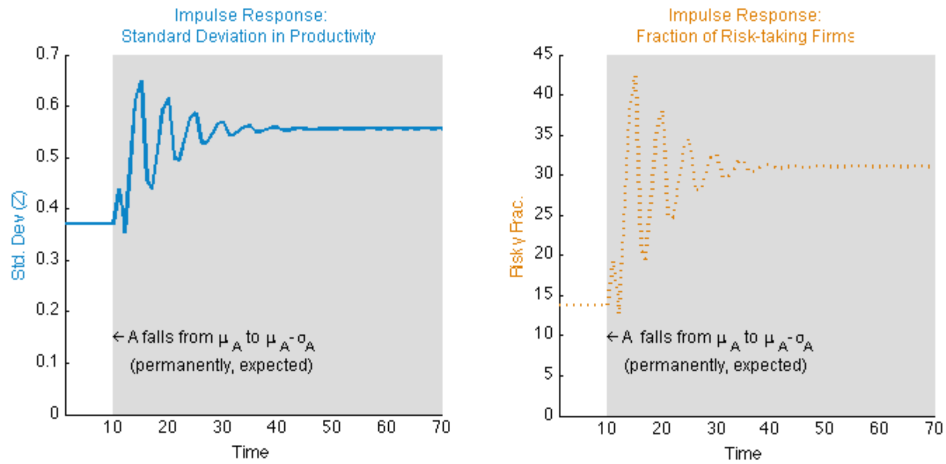

Figure 2.9: Impulse Responses. The figure shows impluse responses to a permanent (and expected) one-standard-deviation drop in the aggregate state. The left panel is the response of cross-sectional productivity dispersion, measured as the standard deviation in realized productivity. The right panel plots the response of the fraction of firms choosing the risky technology.

paths may seem unusual at first glance, but it is the joint work of (1) project choice and (2) capital adjustment costs. Upon the bad shock, as the result of a higher entering threshold, the number of entrants immediately drops to a lower level and then remains constant, and the scale of the economy, measured as the total number of remaining firms, decreases gradually to a new stable level. If capital adjustment costs are shut down, then both the absolute number and the fraction of risk-taking firms jump up upon impulse and drop in the following period. The reason for this sudden jump and drop is that the risky project becomes more appealing to firms with a wider range of capital stock when the shock hits, even though there is a higher probability of bad outcome. Consequently, a large number of firms exit due to their choice of the risky project, which leaves fewer firms remaining in the risky region and this causes the following drop. The absolute number of risky firms then gradually decreases while the fraction increases to a higher level because of the decreasing scale. This up-and-down trend is in line with what is shown in Figure 2.9, which is driven by the project choice. On the other hand, the oscillation is due to the capital adjustment costs, which create firms' inaction in investment and prevent firms from freely changing their capital stocks. Therefore, firms that should be in the risky region in the free adjustment case may now be outside, and vice versa. Note that the fraction of risky firms is around 14% when A is kept at its mean, corresponding to the standard deviation in productivity at about 0.37. Cooper and Haltiwanger (2006) find that the plant-specific idiosyncratic shock has a standard deviation of 0.64. Without assuming idiosyncratic risk, the calibrated stationary version of this model is capable of reproducing at least half of the micro-level standard deviation.

2.4.6 Discussion

2.4.6.1 Firm Size Rotation

The comparison between the model-generated moments and their empirical counterparts suggests that this is not the whole story and that there are some possible extensions for future work. The additional moments in Table 2.6 indicate that the shape of the firm size distribution generated from the model is considerably different from the true one. Without altering the mechanism, introducing further heterogeneity in productivity can at least partly overcome this issue. In addition to that, adding more shocks, such as micro-level idiosyncratic shocks, and allowing for a richer set of risky projects can improve the fit of calibration targets, especially the standard deviation in investment rates. This can also help reduce the extraordinarily high exit rate under aggregate fluctuation. Again, these extensions

Table 2.6: Additional Moments: Transition Rates

5-Year Transition Rates	Model	Data
Small → Exit	0.3491	0.5032
Small → Small	0.2900	0.4203
Small → Large	0.3609	0.0764
Large → Exit	0.2755	0.1803
Large → Small	0.3228	0.0564
Large → Large	0.4017	0.7633
Entry → Small	0.5070	0.7483
Entry → Large	0.4930	0.2517
Data Source: Acemoglu et al. (2011)		

will not alter the mechanism at work.

2.4.6.2 Price Feedback Effect

To avoid computational complexity, the model does not consider the market clearing conditions for either the final good market or the capital market. Apparently, the price feedback effect may erode the quantitative significance of the model. However, with the magnitude reduced, the mechanism remains intact.

In fact, on the final good market, $F(Z,k) = Zk^{\alpha}$ can be interpreted as a firm's profit function, that is, the revenue net of the cost for variable factors, for example, labor and materials. Specifically, assume that a plant faces an inverse demand function $P(y) = By^{-b}$, and therefore its revenue becomes $R(y) = By^{1-b}$. Suppose the actual production function is $y = \tilde{A}k^{\tilde{\alpha}}l^{\tilde{\phi}}$, and the price for other factors is ω. Then after optimization of l, the revenue function becomes

$$R = \left(B\tilde{A}^{1-b}\right)^{1/\left(\tilde{\phi}(1-b)\right)} \left[\tilde{\phi}(1-b)/\omega\right]^{\tilde{\phi}(1-b)/\left(\tilde{\phi}(1-b)-1\right)} k^{\tilde{\alpha}(1-b)/\left(\tilde{\phi}(1-b)-1\right)},$$

and profit function

$$\pi = \left(1 - \tilde{\phi}\,(1-b)\right) R.$$

Redefining variables gives the form of Zk^{α}. Therefore, Z in the model is more appropriately interpreted as measured revenue total factor productivity that includes information from the demand side, instead of the actual production technology. For the same reason, parameter A, shown later in the model, will also be interpreted as the aggregate state of the model economy, and a change in A is more than just a "technology shock." Under this specification, it is easier to link the model to the data because only TFPR (TFP calculated using revenue data) is required for this model, but not TFPQ (actual TFP). Admittedly, TFPR is much easier to compute.

A potentially more interesting extension is to generalize the model in a general equilibrium framework and consider the clearing condition for the capital market. One way to do so is to endogenize the capital market in which exiting firms and shrinking firm can sell their capital holdings to growing ones. In this way, there is an endogenous series of capital prices θ_t, instead of a fixed capital resale price θ. Naturally, θ_t is lower in bad times as more firms reduce their capital stocks, and it is higher in good times as more firms expand. But assuming that firms can employ a one-to-one capital production technique, θ_t will not exceed 1. As a robustness check, I let θ_t be a linear and increasing function of the aggregate state A_t such that $\theta_t = \theta + b_\theta\,(A_t - \mu_A)$ with $b_\theta = 0.5$[13]. The results are presented in Table 2.7. The similarity to the main result is not surprising, because the mechanism remains unchanged.

[13] I also tried $b_\theta = 1, 1.5, 2$ with θ_t constrained to be no higher than 1. The results are very similar.

Table 2.7: Robustness Check: Time Varying θ_t

Cyclicality: Correlations (p-value) with Cyclical Indicators			
		Cyclical Indicators	
Variables of Interests		Avg. Productivity, A	Total Output, O
Productivity Dispersion	$\sigma(Z)$	-0.4154 (0.0000)	-0.7622 (0.0000)
Frac. of Risky Firms	Λ	-0.4296 (0.0000)	-0.6992 (0.0000)
IPR (95%-5%)	IPR_5^{95}	-0.2483 (0.0000)	-0.7424 (0.0000)
Entry Rate	r^{EN}	-0.0581 (0.1943)	-0.8128 (0.0000)
Exit Rate	r^{EX}	-0.4679 (0.0000)	-0.6606 (0.0000)

2.5 Conclusion

Productivity dispersion tends to be larger during recessions. The prevailing view is that increased uncertainty causes a decline in aggregate economic activities. However, although uncertainty matters, this story seems to contradict the observation that recessions lead an increase in productivity dispersion. To complement existing theories, I explore a simple mechanism through which aggregate fluctuations due to standard "technology shocks" can endogenously generate countercyclical dispersion in plant/firm-level productivity. I alter the standard industry dynamics model with business cycle features by incorporating project choice as part of the individual decision problem. By this feature, a firm in this model can then decide the riskiness of its production. The resulting productivity distribution is non-degenerate even if no other heterogeneity is modeled. The model provides the following predictions: small firms are more likely to take risks and have lower survival rates, but conditional on surviving, they exhibit higher productivity; a larger fraction of firms become risky in bad times, which also leads to higher exit rates; and realized micro-level productivity dispersion is larger in recessions.

Chapter 3

Directed Search and Job Rotation

This is paper joint with Fei Li, and it is published on the *Journal of Economic Theory* 148.3 (2013): 1268-1281. [1]

3.1 Introduction

The practice of job rotation is commonly observed in large firms. In the literature, it is well known that a job rotation policy mainly results from learning of the pair-wise match quality between workers and jobs. However, little work has been done to address the impact of job rotation within firms on the labor market. One reason is that the study of job rotation requires a framework that simultaneously considers the internal labor market of a firm and the external labor market. Yet, in the canonical job search model, labor economists' favorite work horse, a firm is treated as a single job vacancy, and therefore it is impossible to distinguish between the internal and external labor market. Recently, many job search papers, including Hawkins (2011), Kaas and Kircher (2011), Lester (2010) and Tan (2012), have shed light on the endogenous determination of firm size, which has the potential to study the interaction between a firm's internal and external labor market.

In this paper, we employ a directed search model with multi-vacancy firms to examine the role of job rotation in the labor market. In particular, we assume that a firm can choose its size by determining the number of job vacancies. A large firm can hire more workers, which requires a higher fixed cost. All workers are ex ante identical, but they may be good at different jobs, which is initially unknown. The match quality between a worker and a job is uncertain when the worker is hired but can be learned afterwards through a process of job rotation. Firms can reassign workers to different positions to partially overcome the loss of mismatch, and larger firms have a higher degree of freedom of reallocation and, therefore, can expect higher revenue per match.

Our main result highlights the impact of job rotation on the labor market. In the unique symmetric equilibrium, we obtain a *positive* correlation between firm size, labor productivity and wage, which is consistent with the stylized facts summarized by Oi and Idson (1999). Without the opportunity of job rotation, however, the correlation between firm size, labor productivity and wage is negative for all parameters, which is the result of a standard directed search model with multi-vacancy firms. In addition, in line with recent empirical findings by Papageorgiou (2011), the model successfully implies a *negative* correlation between firm size and the separation rate.

Our paper is related to the literature in two ways. First, Meyer (1994) and Ortega (2001) point out the learning role of job rotation in firms. They provide a justification for job rotation, but both authors narrow their studies within the boundary of a single firm. As a step further, we apply their insight in a competitive labor market model to study the effect of within-firm job rotation on the external labor market. Papageorgiou (2011) is the only paper that studies the impact of job rotation on the labor market but with a different focus. He pays more attention to the interaction between learning and job reallocation within a firm, while, in contrast, we focus on how the internal labor market in the presence of job rotation affects job allocation in the external labor market. In his model, firm sizes are exogenous rather than endogenously determined as in ours. In our model, the job rotation inside firms (internal) has a feedback effect on firms' contract posting behavior and workers' application behavior (external), observed as variables such as firms' growth rates and size distribution. This feedback is absent in Papageorgiou (2011). In addition, in Papageorgiou's model, the belief of current match quality measuring a worker's performance pins down the wage, which is independent of the firm size once the belief is controlled for. The wage premium in size is then obtained via a comparison of average wages in firms of different sizes. In our model, however, the wage differential in firm size exists conditional on a worker's performance. This provides a testable implication that can distinguish our model from his.

Second, the directed search model we employ follows Montgomery (1991), Peters (1991), Burdett, Shi and Wright (2001), and their later extension by Lester (2010) to the multi-vacancy case. Kaas and Kircher (2011) also study a directed search model with multi-vacancy firms. However, none of these papers can generate a positive relationship between firm size, wage and labor productivity that is in line with observations without introducing *ex ante* exogenously dispersed random productivity;[2] whereas in our model, the presence of learning and job rotation creates an *ex post* heterogeneity among firms and, therefore, can imply a positive relation between labor productivity, wage and firm size. Alternatively, Shi (2002) introduces a frictional product market where large firms pay higher wages to attract workers so that they can produce enough output to fulfill their bigger market share. Tan (2012) allows for local convexity in the production function to generate a positive size-wage differential. Yet, in our model, the production function is concave.

The rest of this paper is organized as follows. We first set up the model and characterize the unique symmetric equilibrium. Next, we derive the implications of our model and discuss the result and compare them to the empirical evidence.

3.2 The Model

3.2.1 Setup

There are N workers and M firms on the market, both of which are ex ante identical. A firm can choose to have multiple vacant positions, each of which requires a worker to form an active match. Denote $\lambda = M/N$ as the firm-worker ratio, which does not represent the labor market tightness due to endogenous vacancy numbers. Following the literature, we first consider the individual decision problem given finite N and M, then we fix λ and take N, M to infinity to approximate the equilibrium in a large labor market.

The quality of a worker-job pair follows a Bernoulli distribution, which is initially unknown upon match but can be perfectly learned later. With probability $\rho \in (0, 1]$,

[2] In both Lester (2010) and Kaas and Kircher (2011), if firms have homogeneous productivity, the relation between wage and firm size is negative.

a match is good and produces 1 unit of output; the match is bad with probability $1-\rho$ and results in 0 output.[3] We assume the match quality is independent across jobs and workers, even within a multi-job firm.

The game has four stages: offer posting stage (I), job searching stage (II), learning and rotation stage (III), and production stage (IV). At Stage I, each firm decides how many vacancies to post, k, and at what wage level, w, where w is potentially a function of k. Firms also announce the firing policy. For simplicity, we assume that they can create $k \in \{1, 2\}$ vacancies with cost $C(k)$, thus the market tightness, defined as the ratio of vacancies to workers, is $\theta \in [\lambda, 2\lambda]$. Without loss of generality, we assume a convex cost function with $C(1) = 0$, $C(2) = C$, $0 < C < \rho$. We assume that wage, $w \in [0, 1]$, does not depend on any further information such as the realized number of applicants and revealed match quality in Stage III. We assume that a firm can commit to the verifiable wage it posts, and the firing strategy, which may depend on the result of learning.[4] Consequently, firms pay the first round of wages to all employees at Stage III and pay the second round only to the remaining ones at Stage IV.

At Stage II, the job searching stage, each worker observes (k, w_k) and the firing rule of every firm and applies to the firms that offer the highest expected payoff. We assume that workers can only apply to a firm, but not to a specific position in

[3]Zero output for a bad match is a strong assumption. However, we assume Bernoulli match qualities to ensure simplicity of the separation rules: separation happens at the zero realization only. In the continuous case, however, the separation follows a cutoff rule such that separation happens for quality below a threshold.

[4]The contract specifies the wage and the firing rule. Without loss of generality, we focus on time-invariant wage contracts. The restricted optimal contract remains optimal in a larger contract space where time-varying wages are permitted because firms and their workers have the same intertemporal rate of substitution of wages. This identical rate of substitution, which is related to the probability of a worker not being fired, stems from the assumptions that (i) there is no discount on the future, (ii) all workers and firms are risk neutral, and (iii) both parties earn zero upon separation.

that firm. If the number of workers that apply for a particular firm exceeds the number of vacancies posted, the firm randomly hires just enough workers; otherwise the firm hires all applicants. Hence, a worker's expected payoff from applying to a firm is determined jointly by both the posted wage and the probability of getting a job.

At Stage III, the learning and rotation stage, a firm randomly assigns hired worker(s) to its position(s) and pays the first round of wage. The firm then learns the match qualities of all job-worker pairs by switching workers to different working positions.[5] In particular, a firm with k jobs and h employees, $1 \leq h \leq k$, learns about the match qualities of all $P_h^k = k!/(k-h)!$ possible worker-job pairs. No production happens at this stage.

At Stage IV, the production stage, a firm is given the option of firing its employee(s) and can reassign remaining ones to specific positions, and then production takes place. An employee gets zero payment once she is fired. After the reallocation of remaining workers, each worker-job pair produces output according to its realized quality. The firm then pays its remaining workers the second round of wage and takes away the rest of total output.

3.2.2 Analysis

The solution concept we adopt is a symmetric rational expectations equilibrium (henceforth, equilibrium), in which each firm chooses to be a large one with the same probability and posts the same contracts, and each worker applies to a large firm with the same probability. The reason for this equilibrium selection is twofold:

[5] We assume that the rotating and learning process serves only to reveal the match qualities but does not generate any production.

first, it delivers a limiting matching technology that has all of the properties required by the competitive model; second, it ensures the nice informational properties of anonymous equilibria in the sense that agents can compute their best replies from aggregate information about the market. We will solve the game backwards. Given any history of Stage II, which will be defined later, a firm learns about the match qualities of all possible worker-job pairs in Stage III, and then, if possible, it assigns jobs to workers to yield the highest revenue. Then we step back to Stage II and characterize the equilibrium in this subgame for any given history in which firms play symmetric strategies. Then, we will characterize each firm's offer posting strategy given the strategies of workers.

Stage IV: Production Stage. At the last stage, firms fire workers when necessary, reallocate remaining ones, and make payments w as promised. It is easy to see that, given any promised $w \geq 0$, the optimal firing happens in one of the two following situations: (1) the worker is unqualified for any position in the firm, or (2) two workers in a large firm are both qualified for one position and not for the other. In the latter case, one worker is enough to produce 1 unit of output and the firm will randomly fire one of the two. The firm then assigns the remaining worker(s), if any, to job(s) in such a way that each worker-job pair is good and produces 1 unit of output.

A small firm fires the employee in case (1) only, so the match is destroyed with probability

$$\phi_{11}(\rho) = 1 - \rho,$$

where $\phi_{kh}(\rho)$ represents the probability of a worker getting fired in a firm with k jobs and h employees; moreover, it is also the separation rate per filled vacancy from the firm's point of view. Hence, with probability $1 - \phi_{11}(\rho)$, the final output is 1 and the worker gets paid w. Similarly, a large firm with one position filled fires the employee also in case (1) only, and the probability of this employee getting fired is

$$\phi_{21}(\rho) = (1-\rho)^2 < \phi_{11}(\rho),$$

which is also associated with 0 earnings to both the employee and the firm. With probability $1-\phi_{21}(\rho)$, 1 unit is produced and the worker gets paid w. Alternatively, workers in a large firm at full capacity may face either of the two cases. From a worker's point of view, case (1) happens with probability $(1-\rho)^2$ and then she is certainly fired, while case (2) means that she is good at only one position and her co-worker has the same ability, so the probability is $2[\rho(1-\rho)]^2$, but she may survive this situation with 1/2 probability. Combined, the overall probability that either one of the two workers will lose her job is

$$\phi_{22}(\rho) = (1-\rho)^2 + \rho^2(1-\rho)^2 < \phi_{11}(\rho).$$

This is also the probability that either one of the two positions will result in 0 output. Obviously, the output is 2 if neither case is realized, 0 if case (1) happens to both workers, and 1 for the rest of the possibilities.

Stage III: Learning and Rotation Stage. A firm learns about the match qualities of all possible worker-job pairs in this stage through the practice of job rotation and pays the promised wage w to employees regardless of the learning results. At the beginning of this stage, the expected output in a firm with k positions and h matched employees, denoted as $F_{kh}(\rho)$, can be computed based on the analysis of

Stage IV, and so can the expected payoff $V_{kh}(\rho; w)$ to its worker(s). In a small firm matched to a worker,

$$\begin{aligned} F_{11}(\rho) &= 1 - \phi_{11}(\rho) = \rho, \\ V_{11}(\rho; w) &= (1 + (1 - \phi_{11}(\rho)))\, w = (1 + \rho)\, w. \end{aligned}$$

The worker receives w at this stage for sure and gets another w if not fired at the next stage, with probability $1 - \phi_{11}(\rho)$. Hence the small firm takes away the rest of the output, $F_{11}(\rho) - V_1(\rho; w)$.

In a large firm, the loss from the mismatch between workers and positions can be partially overcome through job rotation, which results in higher expected labor productivity and a lower separation rate. Specifically, if the firm promises to pay wage w and is only matched to one worker, then

$$\begin{aligned} F_{21}(\rho) &= 1 - \phi_{21}(\rho) = 2\rho - \rho^2 > F_{11}(\rho), \\ V_{21}(\rho; w) &= (1 + (1 - \phi_{21}(\rho)))\, w = \left(1 + 2\rho - \rho^2\right) w. \end{aligned}$$

The firm gets $F_{21}(\rho) - V_{21}(\rho; w)$. If the large firm has two employees and runs at full capacity, each position is expected to produce 1 unit of output with probability $1 - \phi_{22}(\rho)$, so the expected total output from the two positions is:

$$F_{22}(\rho) = 2(1 - \phi_{22}(\rho)) = 4\rho - 4\rho^2 + 4\rho^3 - 2\rho^4 > 2F_{11}(\rho).$$

Observe that $F_{22}(\rho) < 2F_{21}(\rho)$, so the marginal labor productivity in a large firm is decreasing in the number of employees. The payoff to each worker is similar to earlier cases, where both current-stage wage and possible future wage are taken into consideration:

$$V_{22}(\rho; w) = (1 + (1 - \phi_{22}(\rho)))\, w = \left(1 + 2\rho - 2\rho^2 + 2\rho^3 - \rho^4\right) w,$$

and the firm gains an expected profit $F_{22}(\rho) - 2V_{22}(\rho; w)$ now that there are two workers.

Given the ex post incentive compatible separation and job rotation rule, and since there is no strategic interaction at Stages III and IV, matched workers' and firms' payoffs are uniquely pinned down by the contracts they signed. Hence, an equilibrium in our four-stage game is consistent with an equilibrium in a reduced-form two-stage game that includes Stages I and II in the original game, and the payoff is specified as follows: in a firm of k jobs and h workers, a worker's payoff is $V_{kh}(\rho; w)$, and the firm's is $F_{kh}(\rho) - hV_{kh}(\rho; w)$, where $1 \leq h \leq k \leq 2$. In the rest of this paper, we directly solve equilibria of this reduced-form game as those of the whole game.

Stage II: Job Searching Stage. The realization of firms' job posting at Stage I can be summarized by a history vector $H = ((k^j, w^j)_{j=1}^{M})$ listing the number of vacancies and the wages of all M firms. Let $\mathcal{H}$ be the set of all possible H's. In principle, a worker's strategy is defined as $\gamma : \mathcal{H} \rightarrow [0,1]^M$. Given a history H, a worker chooses a vector γ such that (1) γ^j is the probability that he applies to firm $j \in \{1, 2, ..M\}$ and (2) $\sum_{j=1}^{M} \gamma^j = 1$.

Consider the problem of worker i who is deciding whether and to which firm to apply. Firm j posts k^j positions and wage w^j, for $j \in \{1, 2, ..M\}$. If $k^j = 1$, firm j promises its prospective worker the expected payoff $V_1(\rho; w^j)$; if $k^j = 2$, the expected payoff depends on how many workers firm j eventually gets, and it is either $V_{21}(\rho; w^j)$ or $V_{22}(\rho; w^j)$. When the rest $N-1$ workers play identical strategies γ, this worker i chooses strategy $\hat{\gamma}$ to maximize her expected utility

$$\left\{ \begin{array}{c} \sum_{j \text{ s.t. } k^j=1} \hat{\gamma}^j \Omega_1(\gamma^j) V_1(\rho; w^j) \\ + \sum_{j \text{ s.t. } k^j=2} \hat{\gamma}^j \left[\Omega_{21}(\gamma^j) V_{21}(\rho; w^j) + \Omega_{22}(\gamma^j) V_{22}(\rho; w^j)\right] \end{array} \right\} \tag{3.1}$$

where $\Omega_1(\gamma^j)$ stands for the probability that this worker is hired if she applies to firm j which posts $k^j = 1$ positions, that is,

$$\begin{aligned} \Omega_1(\gamma^j) &= \left(1-\gamma^j\right)^{N-1} + \sum_{n=1}^{N-1} \left[\frac{(N-1)!}{n!(N-1-n)!}\right] \left(\gamma^j\right)^n (1-\gamma^j)^{N-1-n} \frac{1}{n+1} \\ &= \frac{1}{N\gamma^j}\left[1-(1-\gamma^j)^N\right], \end{aligned} \tag{3.2}$$

if she is the only applicant, she gets the job for sure; otherwise all applicants get the job with equal probability. The number of applicants at firm j has a binomial distribution. Similarly, $\Omega_{21}(\gamma^j)$ is the probability that this worker is the only applicant at the large firm j and gets a job for sure,

$$\Omega_{21}(\gamma^j) = (1-\gamma^j)^{N-1}, \tag{3.3}$$

and $\Omega_{22}(\gamma^j)$ is the probability that this worker needs to work with someone else in the large firm j,

$$\begin{aligned} \Omega_{22}(\gamma^j) &= \sum_{n=1}^{N-1} \left[\frac{(N-1)!}{n!(N-1-n)!}\right] \left(\gamma^j\right)^n (1-\gamma^j)^{N-1-n} \frac{2}{n+1} \\ &= \frac{2}{N\gamma^j}[1-(1-\gamma^j)^N] - 2(1-\gamma^j)^{N-1}. \end{aligned} \tag{3.4}$$

A symmetric equilibrium at this stage is such that every worker chooses the same application probability vector γ, and moreover, a worker applies to firms of the same size and wage with equal probabilities. Given any history $H = \left((k^j, w^j)_{j=1}^{M} \right)$, $\gamma^* (H)$ is the symmetric solution if $\gamma^* (H)$ is a solution to (3.1) and $\gamma^{*j} (H) = \gamma^{*l} (H)$ if $(k^j, w^j) = \left(k^l, w^l\right)$, $j \neq l$. As mentioned before, we require symmetry across all workers' behavior to ensure an equilibrium that consists of only mixed strategies. In a large market, it is impossible for an individual worker to be fully informed about other workers' job application choices; therefore, modeling it by a mixed-strategy equilibrium is more plausible. More importantly, we assume that a worker applies to firms with identical (k, w) to ensure the anonymity of firms in that workers distinguish between firms only by their sizes and posted wages instead of their names, j. This plays the role of search friction in our model. The symmetry is preserved when we take M and N to infinity.

To model a large market, we will follow the literature and let $M \to \infty$ and $N \to \infty$ such that $\lambda = M/N$ remains constant. Define

$$\mu (k, w) = \lim_{M \to \infty} \left(\sum_{j=1}^{M} 1_{\{(k^j, w^j) = (k, w)\}} \right) /M.$$

At the limit, a history is described by an offer distribution μ. Define the queue length at firm j as $q^j = \lim_{N \to \infty} \gamma^j N$. Using (3.2), (3.3) and (3.4), it is straightforward to establish the hiring probabilities as functions of queue lengths at the limit. If firm j posts one vacancy, then

$$\Omega_1(q^j) = \frac{1}{q^j} \left(1 - e^{-q^j} \right);$$

otherwise, firm j decides to become a large firm and posts two job openings,

$$\begin{aligned}\Omega_{21}(q^j) &= e^{-q^j}, \\ \Omega_{22}(q^j) &= \frac{2}{q^j}\left(1 - e^{-q^j} - q^j e^{-q^j}\right).\end{aligned}$$

In a symmetric equilibrium, given $\mu(k, w)$, all workers play an identical strategy and receive the same and highest utility level denoted as U. Specifically, a worker applies to a small firm j with positive probability only if

$$\Omega_1(q^j)V_1\left(\rho; w^j\right) = U; \tag{3.5}$$

similarly, a worker applies to a large firm j with positive probability only if

$$\Omega_{21}(q^j)V_{21}\left(\rho; w^j\right) + \Omega_{22}(q^j)V_{22}\left(\rho; w^j\right) = U. \tag{3.6}$$

Here, U is referred to as the market utility level in the literature. Solving these two equations gives q^j's as functions of w^j and U. Dropping ρ, define $Q_1(U, w^j)$ as the greater value between the unique q^j as the solution to (3.5) and zero; define $Q_2(U, w^j)$ by doing the same to (3.6). Combined, we have $Q_{k^j}(U, w^j)$, which determines the equilibrium queue length at firm j with (k^j, w^j), when the market utility is U.

Definition 2. *Given an offer distribution $\mu(k^j, w^j)$, a symmetric equilibrium of the Stage II game is characterized by (q^j, U) such that*

1. $q^j = Q_{k^j}(U, w^j)$ *for all* j, *and*

2. $\int Q_{k^j}(U, w^j)\, d\mu(k^j, w^j) = 1/\lambda.$

Hence, workers are indifferent between applying to any firm j as long as $q^j > 0$. At the same time, zero queue length implies that this firm cannot provide the market utility level to workers.

Stage I: Offer Posting Stage. Now take one step back and consider a firm's problem at the limit. Expecting the form of $Q_k(U, w)$ and U, firm j's strategy is to choose a probability distribution μ^j over $\{1, 2\} \times \mathbb{R}^+$, where $\mu^j(k, w)$ is the probability that firm j posts k vacancies and a wage w. If the firm posts a single vacancy, it chooses w_1 to maximize the expected profit,

$$\pi_1^*(U) = \max_{w_1} \{\pi_1(U, w_1) = \Phi_1(Q_1(U, w_1))(F_{11}(\rho) - V_1(\rho; w_1))\}, \tag{3.7}$$

where $\Phi_1(q_1) = q_1\Omega_1(q_1) = 1 - e^{-q_1}$ is the limiting probability that a small firm successfully hires a worker. The market utility level U is taken as given, and the firm can attract applicants only if it can provide U level of expected utility to its potential worker(s). At the same time, the representative firm solves the problem associated with a large one,

$$\pi_2^*(U) = \max_{w_2} \left\{ \pi_2(U, w_2) = \left[\begin{array}{c} \Phi_{21}(Q_2(U, w_2))[F_{21}(\rho) - V_{21}(\rho; w_2)] \\ +\Phi_{22}(Q_2(U, w_2))[F_{22}(\rho) - 2V_{22}(\rho; w_2)] - C \end{array} \right] \right\} \tag{3.8}$$

where $\Phi_{21}(q_2) = q_2\Omega_{21}(q_2) = q_2 e^{-q_2}$ is the probability that a large firm gets only one applicant, and $\Phi_{22}(q_2) = (q_2/2)\Omega_{22}(q_2) = 1 - e^{-q_2} - q_2 e^{-q_2}$ is the probability it gets at least two applicants and therefore two employees. Define

$$\pi^*(U) = \max\{\pi_1^*(U), \pi_2^*(U)\}. \tag{3.9}$$

Naturally, to get the coexistence of both small and large firms, it requires that $\pi^* = \pi_1^* = \pi_2^*$, which is feasible in certain parameter subspaces.

Definition 3. *A symmetric equilibrium of the Stage I game consists of a distribution $\mu^*(k, w)$, a market utility level U^*, and queue lengths q^j, satisfying*

1. $\mu^j(k, w) = \mu^*(k, w)$,
2. $\pi_{k^j}(U^*, w^j) = \pi^*(U^*)$ *if* $d\mu^*(k^j, w^j) > 0$,
3. $\pi_{k^j}(U^*, w^j) \leq \pi^*(U^*)$ *if* $d\mu^*(k^j, w^j) = 0$,
4. (q^j, U^*) *is the equilibrium of the job application game.*

Equilibrium Characterization. In the following proposition, we show that in the unique equilibrium, the only realized history contains identical small firms and/or identical large ones: in a small firm's contract, the proposed wage is w_1^*; in a large firm's contract, it is w_2^*; and the associated equilibrium queue lengths in small and large firms are q_1^* and q_2^*, respectively. Let ϕ^* be the equilibrium probability of becoming a small firm. As a result, the proportion of small firms is $\mu(1, w_1^*) = \phi^*$, and $\mu(2, w_2^*) = 1 - \phi^*$ for large ones. Since workers play a symmetric strategy, they will ignore firms' identity if they proposed the same contract. Hence, we can use σ^* as the probability of applying to the group of small firms, and $1 - \sigma^*$ to the large firms. Immediately, we have

$$\sigma^* = \lambda\phi^* q_1^*, \text{and } 1 - \sigma^* = \lambda(1 - \phi^*) q_2^*,$$

where ϕ^* is the equilibrium probability that a firm becomes a small firm. Given the equilibrium queue lengths q_1^* and q_2^*, (ϕ^*, σ^*) can be uniquely pinned down.

Combining all of the four stages, we can characterize the equilibrium in the following proposition.

Proposition 5. *There exists a list of functions:* $\underline{c}(\rho) \in (0,\rho)$, $\underline{\lambda}(C,\rho) > 0$, *and* $\bar{\lambda}(C,\rho) > 0$. *Fix a set of parameters* (λ, C, ρ) *such that* $C \in (\underline{c}(\rho), \rho)$ *and* $\lambda \in \left(\underline{\lambda}(C,\rho), \bar{\lambda}(C,\rho)\right)$. *There exists a unique symmetric equilibrium in which large firms and small ones coexist. The equilibrium can be characterized by a list of functions* $(\phi^*, w_1^*, w_2^*, \sigma^*)$ *satisfying the following: there exists a unique pair of* (q_1^*, q_2^*), *and a pair of* $(\phi^*, \sigma^*) \in (0,1) \times (0,1)$ *such that*

$$\phi^* = \frac{q_2^* - 1/\lambda}{q_2^* - q_1^*}, \sigma^* = \lambda q_1^* \phi^* = \frac{\lambda q_1^* \left(q_2^* - 1/\lambda\right)}{q_2^* - q_1^*}, q_2^* > q_1^* > 0,$$

and the wages in small and large firm markets are given by

$$\begin{aligned} w_1^* &= \frac{F_{11}(\rho)\, q_1^* e^{-q_1^*}}{\left(1 + F_{11}(\rho)\right)\left(1 - e^{-q_1^*}\right)}, \\ w_2^* &= \frac{F_{21}(\rho) + q_2^* \left[F_{22}(\rho) - F_{21}(\rho)\right]}{1 + F_{21}(\rho) + \left(e^{q_2^*} - 1 - q_2^*\right)\left(F_{22}(\rho) + 2\right)/q_2^*}. \end{aligned}$$

If C, ρ and/or λ lie outside the specified region, which can be decomposed into three regions, there is no heterogeneity in realized firm sizes. The intuition behind these three situations is simple. If $C \in (\underline{c}(\rho), \rho)$ and λ is either too small or too large, firms are also the same size. When λ is too small, there are so few firms in the market relative to workers that it is easy to hire two workers and to take advantage of job rotation. In equilibrium, no firm chooses to become a small one. Similarly, when λ is too large, there are so many firms and vacancies that it is not only costly to post an extra vacancy, but it is also hard to fill both of them in a large firm. In equilibrium, no firm wants to be a large one. The coexistence of small and large firms is only possible when C is high enough compared to ρ, and $\lambda \in \left(\underline{\lambda}(C,\rho), \bar{\lambda}(C,\rho)\right)$. The region in which $C \leq \underline{c}(\rho)$ corresponds to the case of $U^* \geq \rho = F_{11}(\rho)$, and the market utility is so high that a small firm cannot earn a positive profit. As a result, in this region, all firms are the same size. There are two possible cases here: either all firms choose to randomize between being large and not entering by paying an unacceptable wage, or all firms choose to randomize between being small and not entering. The outcome relies on the value of λ. Neither of these two possibilities is of interest. In the following subsection, we focus on the coexistence case and characterize the impact of job rotation on labor market variables.

3.2.3 Implications

In this subsection, we look at the implications of the unique symmetric equilibrium. The model simultaneously gives predictions on relationships between firm size and productivity, separation rate, and wage, which are roughly in line with empirical findings.

Size and Job Rotation Rate. In our model, the job rotation rate is trivially increasing in firm size. We can generalize our model one step further and allow firms to post $1, 2, .., K$ vacancies. Now that a larger firm can overcome the mismatch loss even more via reassignment of jobs, a higher rotation rate will appear. This is consistent with the empirical finding of Papageorgiou (2011). We will see how this higher job rotation benefit of larger firms affects the labor market.

Size and Labor Productivity. The average labor productivity of a small firm is simply $F_{11}(\rho) = \rho$, and that of a large firm is a convex combination $\Phi_{22}F(\rho; 2, 2)/2+$

$\Phi_{21}F(\rho;2,1)$, which is greater than ρ since $F(\rho;2,2) > 2\rho$ and $F(\rho;2,1) > \rho$ for any $\rho \in (0,1)$. As stated before, the marginal labor productivity of a large firm is decreasing in size measured as the number of employees, $F(\rho;2,2) < 2F(\rho;2,1)$, and therefore the production function of a large firm is concave in labor.

Size and Separation Rate. In a recent empirical work, Papageorgiou (2011) analyzes the Survey of Income and Program Participation data and finds that workers in larger firms are less likely to be separated from their firms even conditional on workers' wages. In our paper, for tractability, we assume that after a firm learns the quality of all possible matches between its workers and positions, it has the option to fire incapable employees and create separations. Due to the job rotation feature, large firms have a lower overall separation rate than small firms in our model. In particular, given the specific form of contract, as discussed in the previous section, workers in small firms suffer a separation rate at $\phi_{11}(\rho)$ in Stage IV, and those in large firms working without or with co-workers face the separation rate at $\phi_{21}(\rho)$ or $\phi_{22}(\rho)$. It is obvious that $\phi_{21}(\rho) < \phi_{22}(\rho) < \phi_{11}(\rho)$ for any $\rho \in (0,1)$. Therefore, we have the following result established.

Proposition 6. *The separation rate in a large firm is smaller than that in a small firm.*

Size and Wage Differential. In standard directed search models with multi-vacancy firms, it is well known that small firms always post higher wages in the unique equilibrium.[6] However, this contradicts the observations on the labor market;[7] it is the large firms that pay higher wages to workers. In our model, large

[6] See the discussion in Shi (2002) and Tan (2012).

[7] For example, Brown and Medoff (1989) and Oi and Idson (1999) point out that there exists a positive size-wage differential in the labor market.

firms have the opportunity to reallocate workers over jobs and partially overcome the mismatch between workers and jobs. This job rotation feature creates two simultaneous forces that drive the size-wage differential in different directions. The first effect lies in the increased expected productivity of large firms. When their expected productivity is higher, large firms may be able and willing to pay higher wages to their workers, which makes their job offers more attractive to workers. The second effect is due to the reduced job separation rate in large firms. Lower unemployment risk in large firms works together with the first effect to pull up the expected utility that large firms promise to their applicants, that is, $V_2 = (\Omega_{21}V_{21} + \Omega_{22}V_{22}) / (\Omega_{21} + \Omega_{22}) > V_1$. However, the smaller separation rate can potentially push wages down. Taking both effects into consideration, we claim that, when the mismatch risk is high compared to the extra cost of becoming a large firm, large firms can provide higher promised utility; and when the mismatch risk is even higher so that the first effect dominates, large firms pay higher wages.

Result 2. *Large firms offer lower wages than small firms if there is no mismatch, $\rho = 1$. For any $\rho \in (0,1)$, there exists a $\bar{c}(\rho) \in (\underline{c}(\rho), \rho]$ such that for any $C \in (\underline{c}(\rho), \bar{c}(\rho))$, $V_2 > V_1$. Furthermore, when ρ and C are small enough, there exist a set of (ρ, C) such that $w_2^* > w_1^*$.*

We provide a numerical illustration of this result due to the difficult derivation of an analytical proof. In Figure 3.1, we illustrate how w_1/w_2 and V_1/V_2 depend on C and ρ. When $\rho = 1$, we replicate the result of a standard directed search model with multi-vacancy firms, simply because there is no risk of mismatch. In this case, large firms offer lower wages for any positive C. When ρ is small, it is possible to obtain the wage premium of large firms. The intuition is as follows.

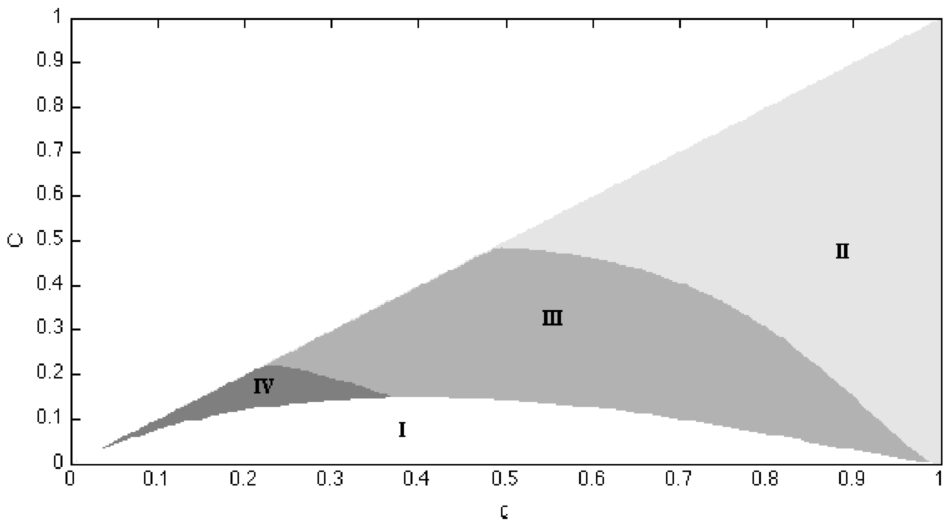

Figure 3.1: Decomposition of (ρ, C)-space. I: $C \leq \underline{c}(\rho)$, no co-existence of firms of two sizes. II $V_1 > V_2$ and $w_1 > w_2$. III $V_1 < V_2$ and $w_1 > w_2$. IV $V_1 < V_2$ and $w_1 < w_2$.

Smaller ρ implies a higher probability of mismatch and, consequently, a greater job rotation benefit and a higher wage premium; thus the wage premium is decreasing in ρ. There are four relevant regions. Region I corresponds to the case of $C \leq \underline{c}(\rho)$, which is not of interest. In region II, C is relatively high so becoming a large firm is costly, and ρ is large and the advantage of rotation is limited; thus, small firms provide more promising offers in the equilibrium, $V_1 > V_2$. In region III, (ρ, C) is moderate and the advantage of rotation raises large firms' expected productivity so that their offer becomes more attractive than those of small firms, and $V_2 > V_1$. However, since workers in large firms face a smaller unemployment risk, when (ρ, C) belongs to this region, to provide higher expected utility, large firms do not need to pay high wages, so $w_2^* < w_1^*$. In region IV, (ρ, C) is small enough, and the difference in unemployment risk is limited, hence $w_2^* > w_1^*$.

Decomposition of (ρ, C)-space. I: $C \leq \underline{c}(\rho)$, no co-existence of firms of two sizes. II $V_1 > V_2$ and $w_1 > w_2$. III $V_1 < V_2$ and $w_1 > w_2$. IV $V_1 < V_2$ and $w_1 < w_2$.

For standard directed search models to generate a positive correlation between firm size and wage, an exogenous productivity difference is required. In particular, Kaas and Kircher (2011) and Lester (2010) assume that firms randomly draw their productivity levels from a pre-determined distribution before they enter the labor market, and high productivity firms decide to be large and low productivity firms choose otherwise. If the ex ante distribution of productivity is dispersed enough, this technology difference can overcome the frictional effect of coordination failure and can generate a reasonable size-wage differential. In their models, large firm size and a wage premium are the consequence of high productivity. Our model suggests a somewhat reversed direction of such a relationship: even with ex ante homogeneity assumed, large firms may emerge, taking advantage of the opportunity of job rotation, which in turn induces high productivity and a wage premium.

3.3 Vacancy Yield and Informative Interview

In this section, we study the vacancy yield[8]. In our baseline model, we assume vacancies are ex ante homogeneous across firms. Let v_k be the equilibrium vacancy yield of firms posting k vacancies in our benchmark model, which is the probability of filling a position in these firms, then we have $v_1 = \Phi_1(q_1^*)$. In a large firm, it is straightforward to see that $\Phi_{21}(q_2^*) = 2v_2(1-v_2)$ and $\Phi_{22}(q_2^*) = (v_2)^2$, so $v_2 = \Phi_{22}(q_2^*) + \Phi_{21}(q_2^*)/2$. Our simulation shows $v_1 < v_2$ for any $\rho \in (0,1]$ and $C \in (\underline{c}(\rho), \rho)$, but it is inconsistent with the empirical relation between vacancy yield and firm size, which is negative. This inconsistency is a typical result in

[8] We thank an associate editor for encouraging us to investigate this issue in our framework.

directed search models, for example, Lester (2010), because wages play an allocative role in the workers' application decision. Nonetheless, in comparison to a model without the opportunity of job rotation, $\rho = 1$, our model here predicts a greater disparity between the vacancy yields of firms with one vacancy and those with multiple vacancies, i.e., the difference between v_2 and v_1 is amplified as ρ becomes smaller.

An important factor, however, is missing in our main model, as well as in most directed search models. As argued by Davis, Faberman, and Haltiwanger(2010), firms of different sizes have heterogeneous job recruiting standards due to the pre-existing heterogeneity on both sides of the labor market. Acknowledging this, we now extend our main model to investigate the possibility that large firms have a different job recruiting standard from small firms.

Suppose a large firm, by paying the extra cost C, can afford a more sophisticated human resources department and, therefore, can draw an informative but noisy signal about the match quality between potential employees and their positions.[9] We introduce a heterogeneity of interview technology among firms of different sizes to capture the idea that large firms have higher job recruiting standards than small firms. To simplify the analysis, we focus on the following signal-generating technology. If a worker is good at neither position, a bad signal is realized with probability $1-\delta$, where $\delta \in (0,1)$.[10] Hence, conditional on being matched with a large firm, the probability that a worker passes the interview is $\eta = 1 - (1-\rho)^2 (1-\delta)$ which is

[9]We fold the cost of the additional screening technology into the second vacancy posting cost. Hence, large firms are equipped with this technology automatically. It is, however, possible to endogenize this decision; see, for example, Galenianos (2012).

[10]We assume that large firms cannot acquire workers' match quality information position by position, which implies that firms will randomly allocate a qualified employee over positions. Since our interest is not in studying the effect of interviews on firms' job assignment to new workers, we believe this assumption does not lose any generality.

close to zero when $\rho, \delta \to 0$. If a worker passes the interview, his posterior of being good at each position is given by $\bar{\rho} = \frac{\rho}{\rho+(1-\rho)(\delta+\rho(1-\delta))} \in (\rho, 1)$. Similar analysis yields the equilibrium wages w_1^* in small firms and $w_2^*(\bar{\rho})$[11] in large ones, and vacancy yields in small and large firms are given by $v_1 = \Phi_1$ and $v_2 = \eta(\Phi_{22} + \Phi_{21}/2)$. When δ is small (the signal is precise), large firms are very selective, and therefore, the vacancy yield in large firms can be smaller than that in small firms. Figure 2 shows some numerical examples. For small ρ and C, when δ is small, $v_1 < v_2$, and $w_1^* < w_2^*$. Since a match is good with probability $\bar{\rho} > \rho$ in large firms, both the productivity difference and the separation rate difference between large firms and small firms are amplified. On the other hand, the interview effect will decrease the possibility of job rotation. However, in our model, since the job rotation rate in small firms is always zero, our prediction on the relation between job rotation rate and firm size still holds.

We assume that large firms can only draw signals from matched workers. What if they could draw signals from all applicants? The result will not change qualitatively. The reason is as follows. In equilibrium, a large firm faces finitely many applicants. Even though there are more than 2 applicants, the probability that the firm cannot hire enough workers is always positive if $\delta \in (0, 1)$. When both ρ and δ are small, the vacancy yield can be arbitrarily small. Hence, our prediction on the relation between vacancy yield and firm size still holds.

[11] The wage in large firms, $w_2^*(\bar{\rho})$ is obtained by replacing ρ by $\bar{\rho}$ in the expression of w_2^* in Proposition 1.

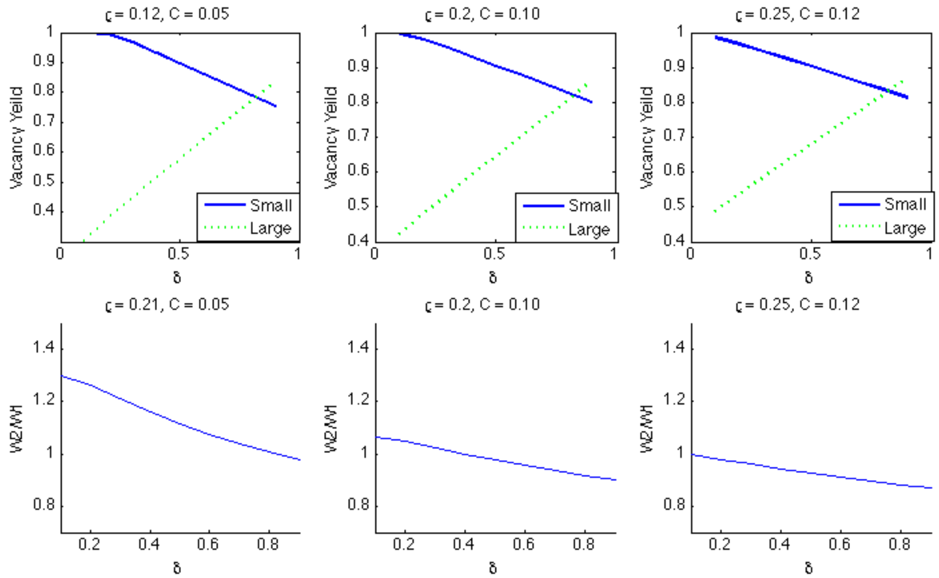

Figure 3.2: Vacancy Yield and Wage Ratio at Different ρ, C, and δ.

3.4 Conclusion

We modified a standard directed search model to explain the size-wage differential observed in the labor market, highlighting the effect of the practice of job rotation. However, in contrast to the standard directed search model with multi-vacancy firms, our modified model can generate a positive correlation between firm size and wage without introducing any ex ante exogenous productivity heterogeneity or imposing any non-concave production function assumptions. We assume ex ante homogeneous firms and workers and initially unknown match quality that determines labor productivity. Firm sizes are endogenously determined. By paying an extra cost, a large firm benefits from the opportunity to rotate workers so as to partially overcome the loss of mismatch. As a result, in the unique symmetric equilibrium, large firms have higher labor productivity and wages, and a lower separation rate. In future research, we would like to study the interaction between internal labor markets and external labor markets in a fully dynamic model.

Appendix A

Appendix for Chapter 1

A.1 Appendix: Omitted Proofs

This part of the appendix contains the proofs that are omitted from the main context.

A.1.1 Derivation of Price Indices

Household's FOC

$$\beta^t \eta_t \theta_{it} = \lambda_t p_{it} c_{it.}$$

So

$$\frac{p_{it} c_{it}}{\theta_{it}} = \sum p_{it} c_{it}$$

since $\sum \theta_{it} = 1$,

$$\sum p_{it} c_{it} = \prod \left(\sum p_{it} c_{it} \right)^{\theta_{it}} = \prod \left(\frac{p_{it} c_{it}}{\theta_{it}} \right)^{\theta_{it}} = P_t^C C_t.$$

Plugging the indices back into the household's FOC's yields equations (1.4), (1.5), and (1.6).

Similarly, firm i's FOC:

$$p_{jt}x_{ijt} = \alpha_i \omega_{ij} E_t \left(\Lambda_{t,t+1} s_{i,t+1}\right)$$

so

$$\frac{p_{jt}x_{ijt}}{\omega_{ij}} = \sum_j p_{jt}x_{ijt}$$

and since $\sum \omega_{ij} = 1$

$$\sum_j p_{jt}x_{ijt} = \prod \left(\sum_j p_{jt}x_{ijt}\right)^{\omega_{ij}} = \prod \left(\frac{p_{jt}x_{ijt}}{\omega_{ij}}\right)^{\omega_{ij}} = P^X_{it} X_{it}.$$

A.1.2 Proof of Proposition 1

Proof. Firm i's FOC for labor demand

$$w_{it}h_{it} = (1-\alpha_i)\, s_{it}.$$

HH's budget constraint

$$\sum p_{it}c_{it} = \sum (1-\alpha_i)\, s_{it} + \sum \pi_{it}.$$

Market clearing

$$p_{it}c_{it} + \sum_j p_{it}x_{jit} = \theta_{it} P^C_t C_t + \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt} P^C_t C_t = s_{it}$$

So

$$\begin{aligned}
\frac{s_{it}}{P_t^C C_t} &= \theta_{it} + \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt} \\
\frac{\sum s_{it}}{P_t^C C_t} &= 1 + \sum_j \alpha_j \tilde{f}_{jt}.
\end{aligned}$$

□

Note that the proof also implies

$$\begin{aligned}
\frac{c_{it}}{y_{it}} &= \frac{\theta_{it}}{\theta_{it} + \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt}} \\
\frac{x_{jit}}{y_{it}} &= \frac{\alpha_j \omega_{ji} \tilde{f}_{jt}}{\theta_{it} + \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt}}.
\end{aligned}$$

and HH labor supply condition

$$\begin{aligned}
h_{it}^{\varepsilon} &= \eta_t \frac{w_{it}}{P_t^C C_t} \\
h_{it}^{1+\varepsilon} &= \eta_t \frac{(1-\alpha_i) s_{it}}{P_t^C C_t} = \eta_t (1-\alpha_i) \left(\theta_{it} + \sum_j \alpha_j \omega_{ji} \tilde{f}_{jt} \right).
\end{aligned}$$

A.2 Appendix: Data and Estimation

A.2.1 Data Description and Transformation

The U.S. data used in this paper are from the Industry Economic Accounts of the Bureau of Economic Analysis. Specifically, I take the Use tables and Make tables (before redefinition) between 1997 and 2012, also the GDP by industry tables of the same time. I exploit the information on 65 private industries in each table.

Industries are defined according to the 2007 NAICS, roughly at the three digit level. The detailed information on these industries is included in subsection A.2.3.

The GDP by industry (value added) tables contain value added, gross output, intermediate inputs expenditure by industry data, as well as corresponding price indices, for 69 industries (65 private and 4 non-private). The Make table of a given year documents the production of each commodity by each industry, measured in millions of dollars. A Make table is a 69 industries by 71 commodities matrix ($Make$). The Use table of a given year records the expenditure made by each industry on each commodity used as input, measured in millions of dollars, while it also records the final demand of each commodity. A Use table consists of a 71 commodities by 69 industries matrix (Use) and 71 commodities by 11 groups of final users matrix.

To get consistent industry-by-industry data on value added, gross output, detailed input uses, and final demand for 65 private industries, the Make table of year 2007 is used to transform the commodity-by-industry/final-user Use tables into industry-by-industry/final-user tables. Let $Make_{2007}$ be the original 2007 Make table and define $Make^{I}_{2007}$ such that each (i, j) element is given by

$$\begin{aligned} Make^{I}_{2007}(i,j) &= \frac{Make_{2007}(i,j)}{\sum_k Make_{2007}(k,j)} \\ &= \frac{\text{Industry } i\text{'s production of Commodity } j}{\text{Total output of Commodity } j}. \end{aligned}$$

Therefore each column of $Make^{I}_{2007}$ sums up to 1. The Use table of the year 2007 are used to calibrate the production technoloy parameters, Ω. Let Use_{2007} be the origianl 2007 Use table (71 commodities by 69 industries) and define Use^{IxI}_{2007} such that

$$Use^{IxI}_{2007} = Make^{I}_{2007} \times Use_{2007},$$

then each element $Use^{IxI}_{2007}(i,j)$ is industry j's use of industry i's output. Then each (i,j) element of Ω is simply

$$\Omega(i,j) = \frac{Use^{IxI,\text{ Private}}_{2007}(j,i)}{\sum_k Use^{IxI,\text{ Private}}_{2007}(k,i)}$$

where $Use^{IxI,\text{ Private}}_{2007}$ is the transformed Use table for the private industries, first 65-by-65 block of Use^{IxI}_{2007}.

For each year, the final demand for each commodity is a 71-by-1 vector, calculated as the total final use for each commodity adding back the imports. Adjusting each vector with $Make^{I}_{2007}$ yields the final demand for each industry's output in every year, the first 65 elements of which are the private sector. Normalizing each year's total final demand to be 1 gives the consumption shares, θ_t, from 1997 to 2012. Similarly, θ is obtained averaging the final demand over the sample period.

A.2.2 Estimation of (κ, N, T)

Denote the unconditional variance-covariance matrix of θ_t as $\Sigma^{\theta}(\kappa)$, such that:

$$\begin{aligned}
\Sigma^{\theta}(\kappa) &= Var(\theta_t;\kappa), \\
\Sigma^{\theta}_{ii}(\kappa) &= Var(\theta_{it};\kappa) = \frac{\theta_i(1-\theta_i)}{\kappa+1}, \forall i \\
\Sigma^{\theta}_{ij}(\kappa) &= Cov(\theta_{it},\theta_{jt};\kappa) = -\frac{\theta_i\theta_j}{\kappa+1}, \forall i \neq j.
\end{aligned}$$

By the independence assumption,

$$\Sigma^{\Delta\theta}(\kappa) = Var(\Delta\theta_t) = Var(\theta_t - \theta_{t-1}) = 2\Sigma^{\theta}(\kappa).$$

The news about θ_t received in period s, $\forall s < t$, is independently drawn from the multinomial distribution wiht N trials,

$$\mathbf{m}_t^s := (m_{1t}^s, ..., m_{nt}^s)' \sim Multinomial(N, \theta_t).$$

Conditional on θ_t, the moments of the news satisfies

$$\begin{aligned} E(\mathbf{m}_t^s|\theta_t) &= N\theta_t \\ Var(m_{it}^s|\theta_t) &= N\theta_{it}(1-\theta_{it}), \forall i, \forall s \\ Cov\left(m_{it}^s, m_{jt}^s|\theta_t\right) &= -N\theta_{it}\theta_{jt}, \forall i \neq j, \forall s. \end{aligned}$$

Therefore, the unconditional variance-covariance matrix of the news, denoted as Σ^m, is such that

$$\begin{aligned} \Sigma^m(\kappa, N) &= E(E(\mathbf{m}_t^s\mathbf{m}_t^{s\prime}|\theta_t)) - E(E(\mathbf{m}_t^s|\theta_t))\,E(E(\mathbf{m}_t^{s\prime}|\theta_t)) \\ &= N(N+\kappa)\,\Sigma^{\theta}(\kappa), \forall t, \forall s < t. \end{aligned}$$

Note that, for each piece of news, the unconditional variance-covariance matrix $\Sigma^m(\kappa, N)$ is the same, namely, $\Sigma^m(\kappa, N)$ depends neither on the target time t nor on the news arrival time $s < t$. In fact, $\Sigma^m(\kappa, N)$ remains the same for any T.

For a given T, the forecast vector at each time t can be written as

$$\tilde{\mathbf{f}}_t = Const + \sum_{\tau=1}^{T} \beta^{\tau-1} \left(\Omega' \mathbf{A}\right)^{\tau-1} \frac{\sum_{\tau'=0}^{T-\tau} \mathbf{m}_{t+\tau}^{t-\tau'}}{\kappa + (T+1-\tau) N},$$

where $Const$ is a time-invariant constant vector. Consequently, due to the independence assumption on the news, the unconditional variance-covariance matrix of $\tilde{\mathbf{f}}_t$, $\Sigma^f(\kappa, N, T)$, has the following form

$$\begin{aligned}\Sigma^f(\kappa, N, T) &= Var\left(\tilde{\mathbf{f}}_t; \kappa, N, T\right) \\ &= N(N+\kappa) \sum_{\tau=1}^{T} \frac{\beta^{2(\tau-1)} (T-\tau+1)}{(\kappa + (T+1-\tau) N)^2} \left(\Omega' \mathbf{A}\right)^{\tau-1} \Sigma^\theta(\kappa) \left(\mathbf{A}\Omega\right)^{\tau-1}.\end{aligned}$$

Now I look at the observable variables. Denote the input use by industry i at time t as u_{it}, $u_{it} = P_{it}^X X_{it}$, and $\mathbf{u}_t = (u_{1t}, ..., u_{nt})'$. Hence industry i's value added at t is $v_{it} := s_{it} - u_{it}$, $\mathbf{v}_t = (v_{1t}, ..., v_{nt})'$, and the aggregate value added, namely GDP, is $V_t = P_t^C C_t = \sum_i v_{it}$. By Proposition 1, we have the following ratios,

$$\begin{aligned}\frac{\mathbf{s}_t}{V_t} &= \theta_t + \beta\Omega' A \tilde{\mathbf{f}}_t = \theta_t + \Omega' \frac{\mathbf{u}_t}{V_t} \\ \frac{\mathbf{u}_t}{V_t} &= \beta A \tilde{\mathbf{f}}_t \\ \frac{\mathbf{v}_t}{V_t} &= \theta_t + \beta\Omega' A \tilde{\mathbf{f}}_t - \beta A \tilde{\mathbf{f}}_t = \theta_t + \beta\left(\Omega' - I\right) A \tilde{\mathbf{f}}_t.\end{aligned}$$

Table A1. Goodness of Fit

		Data ($\times 10^{-4}$)	Estimated ($\times 10^{-4}$)
Step 1	$\overline{\sigma_i^\theta}$	5.29	8.44
	$std\left(\sigma_i^\theta\right)$	6.59	5.07
Step 2	$\overline{\sigma^{u/V}}$	12.76	3.36
	$std\left(\sigma^{u/V}\right)$	13.76	2.11
	$\overline{\sigma^{s/V}}$	14.74	11.89
	$std\left(\sigma^{s/V}\right)$	15.91	7.08
	$\overline{\sigma^{v/V}}$	8.13	13.19
	$std\left(\sigma^{v/V}\right)$	7.63	7.77

The variation in each variable over time is then

$$\begin{aligned}
\Sigma^{s/V}(\kappa, N, T) &= Var\left(\frac{\mathbf{s}_t}{V_t}\right) = \Sigma^\theta + \beta^2 \Omega' A \Sigma^f A \Omega \\
\Sigma^{u/V}(\kappa, N, T) &= Var\left(\frac{\mathbf{u}_t}{V_t}\right) = \beta^2 A \Sigma^f A \\
\Sigma^{v/V}(\kappa, N, T) &= Var\left(\frac{\mathbf{v}_t}{V_t}\right) = \Sigma^\theta + \beta^2 \left(\Omega' - I\right) A \Sigma^f A \left(\Omega - I\right).
\end{aligned}$$

All the variables, $\mathbf{s}_t/V_t$, $\mathbf{u}_t/V_t$, and $\mathbf{v}_t/V_t$, can be directly calculated from data, so are the variances. However, in order to eliminate the time trend, I use the changes instead: $\Delta\mathbf{s}_t/V_t = \mathbf{s}_t/V_t - \mathbf{s}_{t-1}/V_{t-1}$, $\Delta\mathbf{u}_t/V_t$, and $\Delta\mathbf{v}_t/V_t$ similarly defined.

Starting from $\Sigma^{\Delta u/V}$,

$$\Sigma^{\Delta u/V} = \beta^2 A \Sigma^{\Delta f} A$$

where

$$\begin{aligned}
\Sigma^{\Delta f} &= Var\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{f}}_{t-1}\right) \\
&= 2\Sigma^f - Cov\left(\tilde{\mathbf{f}}_t, \tilde{\mathbf{f}}_{t-1}\right) - Cov\left(\tilde{\mathbf{f}}_{t-1}, \tilde{\mathbf{f}}_t\right) \\
&= 2\Sigma^f - \Gamma^f - \left(\Gamma^f\right)'
\end{aligned}$$

with $\Gamma^f = Cov\left(\tilde{\mathbf{f}}_t, \tilde{\mathbf{f}}_{t-1}\right) = \left[Cov\left(\tilde{\mathbf{f}}_{t-1}, \tilde{\mathbf{f}}_t\right)\right]'$ being the $(t, t-1)$ covariance matrix. Note that, when $T = 1$, $\tilde{\mathbf{f}}_t$ and $\tilde{\mathbf{f}}_{t-1}$ are independent, so $\Sigma^{\Delta f}(T = 1) = 2\Sigma^f$. For other T,

$$\begin{aligned}
\Gamma^f &= Cov\left(\tilde{\mathbf{f}}_t, \tilde{\mathbf{f}}_{t-1}\right) \\
&= \frac{T-1}{(\kappa + TN)(\kappa + (T-1)N)}\Sigma^m \beta A\Omega \\
&\quad + \frac{T-2}{(\kappa + (T-1)N)(\kappa + (T-2)N)}\beta\Omega' A\Sigma^m (\beta A\Omega)^2 \\
&\quad + ... + \frac{1}{(\kappa + 2N)(\kappa + N)}(\beta\Omega' A)^{T-2}\Sigma^m (\beta A\Omega)^{T-1}.
\end{aligned}$$

Next, the other two variances.

$$\begin{aligned}
\Sigma^{\Delta s/V} &= Var\left(\theta_t - \theta_{t-1} + \beta\Omega' A\tilde{\mathbf{f}}_t - \beta\Omega' A\tilde{\mathbf{f}}_{t-1}\right) \\
&= Var(\Delta\theta_t) + Var\left(\beta\Omega' A\Delta\tilde{\mathbf{f}}_t\right) \\
&- Cov\left(\theta_t, \beta\Omega' A\tilde{\mathbf{f}}_{t-1}\right) - Cov\left(\beta\Omega' A\tilde{\mathbf{f}}_{t-1}, \theta_t\right) \\
&= 2\Sigma^\theta + \beta^2\Omega' A\Sigma^{\Delta f} A\Omega - \beta\Gamma^{\theta,f} A\Omega - \beta\Omega' A\left(\Gamma^{\theta,f}\right)' \\
\Sigma^{\Delta v/V} &= Var\left(\theta_t - \theta_{t-1} + \beta(\Omega' - I)A\left(\tilde{\mathbf{f}}_t - \tilde{\mathbf{f}}_{t-1}\right)\right) \\
&= 2\Sigma^\theta + \beta^2(\Omega' - I)A\Sigma^{\Delta f} A(\Omega - I) - \beta\Gamma^{\theta,f} A(\Omega - I) \\
&+ \beta(\Omega' - I)A\left(\Gamma^{\theta,f}\right)'
\end{aligned}$$

where $\Gamma^{\theta,f} = Cov\left(\theta_t, \tilde{\mathbf{f}}_{t-1}\right)$ is the covariance matrix between θ_t and $\tilde{\mathbf{f}}_{t-1}$,

$$\begin{aligned}
\Gamma^{\theta,f} &= Cov\left(\theta_t, \tilde{\mathbf{f}}_{t-1}\right) \\
&= Cov\left(\theta_t, \frac{\sum_{s=1}^{T} \mathbf{m}_{t-s}^{t}}{\kappa + TN}\right) = \frac{T}{\kappa + TN} Cov\left(\theta_t, \mathbf{m}_{t-s}^{t}\right) \\
&= \frac{TN}{\kappa + TN} \Sigma^{\theta} = \left(\Gamma^{\theta,f}\right)'.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Sigma^{\Delta s/V} &= 2\Sigma^{\theta} + \beta^2 \Omega' A \Sigma^{\Delta f} A \Omega - \frac{\beta TN}{\kappa + TN}\left(\Sigma^{\theta} A \Omega + \Omega' A \Sigma^{\theta}\right) \\
\Sigma^{\Delta v/V} &= 2\Sigma^{\theta} + \beta^2 \left(\Omega' - I\right) A \Sigma^{\Delta f} A \left(\Omega - I\right) \\
&- \frac{\beta TN}{\kappa + TN}\left(\Sigma^{\theta} A \left(\Omega - I\right) + \left(\Omega' - I\right) A \Sigma^{\theta}\right).
\end{aligned}$$

The data counterparts of the unconditional variance-covariance matrices $\Sigma^{\Delta s/V}$, $\Sigma^{\Delta u/V}$, $\Sigma^{\Delta v/V}$ can be directly calculated. Similarly, the unconditional variance-covariance matrix of changes in the product-specific preference $\Sigma^{\Delta\theta}$ can also be calculated using the realized θ_t. The estimation strategy consists of two steps and picks κ and (N, T) sequentially. Step one picks κ to minimize the distance between model and data variables $\Sigma^{\Delta\theta}$, specifically,

$$\hat{\kappa} = \arg\min_{\kappa} \sum_{i=1}^{n} \left(\Sigma_{i,i}^{\Delta\theta,Model}\left(\kappa\right) - \Sigma_{i,i}^{\Delta\theta,Data}\right)^2$$

where $\Sigma_{i,i}^{\Delta\theta} = Var\left(\Delta\theta_{it}\right)$, the i-th element on the diagnal. And the second step finds (N, T) in an analogous way with $\hat{\kappa}$ given,

$$\left(\hat{N}, \hat{T}\right) = \arg\min_{N,T} \sum_{i=1}^{n} \left[\begin{array}{c} \left(\Sigma_{i,i}^{\Delta s/V,Model}\left(N,T;\hat{\kappa}\right) - \Sigma_{i,i}^{\Delta s/V,Data}\right)^2 \\ + \left(\Sigma_{i,i}^{\Delta u/V,Model}\left(N,T;\hat{\kappa}\right) - \Sigma_{i,i}^{\Delta u/V,Data}\right)^2 \\ + \left(\Sigma_{i,i}^{\Delta v/V,Model}\left(N,T;\hat{\kappa}\right) - \Sigma_{i,i}^{\Delta v/V,Data}\right)^2 \end{array} \right].$$

The results, as shown in Table 2, are $\kappa = 15300$, $N = 9100$, and $T = 1$. Table A1 summarizes the goodness of fit.

A.2.3 NAICS Code and Industry Description

The 2007 NAICS code and description of each of the 65 private industries are listed in the following table.

NAICS Code	Industry Description
111CA	Farms
113FF	Forestry, fishing, and related activities
211	Oil and gas extraction
212	Mining, except oil and gas
213	Support activities for mining
22	Utilities
23	Construction
321	Wood products
327	Nonmetallic mineral products
331	Primary metals
332	Fabricated metal products
333	Machinery

NAICS Code	Industry Description
334	Computer and electronic products
335	Electrical equipment, appliances, and components
3361MV	Motor vehicles, bodies and trailers, and parts
3364OT	Other transportation equipment
337	Furniture and related products
339	Miscellaneous manufacturing
311FT	Food and beverage and tobacco products
313TT	Textile mills and textile product mills
315AL	Apparel and leather and allied products
322	Paper products
323	Printing and related support activities
324	Petroleum and coal products
325	Chemical products
326	Plastics and rubber products
42	Wholesale trade
441	Motor Vehicle and Parts Dealers
445	Food and Beverage Stores
452	General Merchandise Stores
4A0	Other Retail
481	Air transportation
482	Rail transportation
483	Water transportation
484	Truck transportation

NAICS Code	Industry Description
485	Transit and ground passenger transportation
486	Pipeline transportation
487OS	Other transportation and support activities
493	Warehousing and storage
511	Publishing industries (includes software)
512	Motion picture and sound recording industries
513	Broadcasting and telecommunications
514	Information and data processing services
521CI	Federal Reserve banks, credit intermediation, and related activities
523	Securities, commodity contracts, and investments
524	Insurance carriers and related activities
525	Funds, trusts, and other financial vehicles
531	Real estate
532RL	Rental and leasing services and lessors of intangible assets
5411	Legal services
5415	Computer systems design and related services
5412OP	Miscellaneous professional, scientific, and technical services
55	Management of companies and enterprises
561	Administrative and support services
562	Waste management and remediation services
61	Educational services
621	Ambulatory health care services

NAICS Code	Industry Description
622	Hospitals
623	Nursing and residential care facilities
624	Social assistance
711AS	Performing arts, spectator sports, museums, and related activities
713	Amusements, gambling, and recreation industries
721	Accommodation
722	Food services and drinking places
81	Other services, except government

Appendix B

Appendix for Chapter 2

Proof of Proposition 3

Let k_0^A be such that $A\left(k_0^A\right)^\alpha = k^*$, where $k^* = (\alpha\beta A)^{1/(1-\alpha)}$, then,

$$k_0^A = (\alpha\beta)^{\frac{1}{\alpha(1-\alpha)}} A^{\frac{1}{1-\alpha}} < k^*.$$

Similarly, let $k_0^{\bar{z}}$ be such that $\bar{z}\,(k_0^{\bar{z}})^\alpha = k^*$, that is

$$k_0^{\bar{z}} = (\alpha\beta p)^{\frac{1}{\alpha(1-\alpha)}} \bar{z}^{\frac{1}{1-\alpha}} < k_0^A.$$

When $k_0 < k_0^{\bar{z}}$, the firm that stays in period 1 will further invest all w_1 in a safe project since $w_1 < k^*$. Let k_0^{II} be the investment level at which a firm is indifferent between investing in a safe project and a risky one, meaning $V_1^1\left(Ak_0^{II\alpha}\right) = pV_1^1\left(\bar{z}k_0^{II\alpha}\right) + (1-p)\,V^0$. When Condition 1 holds, it is straightforward to show that

$$k_0^{II} = \left[\frac{(1-p)\, V^0}{\beta \bar{z}^{1+\alpha} \left(p^{1+\alpha} - p^2\right)} \right]^{\frac{1}{\alpha^2}} < k_0^{\bar{z}}.$$

Let k_0^I be the investment level at which a firm is indifferent between exiting and investing in a risky project, that is, $V^0 = pV_1^1 \left(\bar{z} k_0^{I\alpha}\right) + (1-p)\, V^0$, then

$$k_0^I = \left[\frac{V^0}{\beta A \bar{z}^{\alpha}} \right]^{\frac{1}{\alpha^2}} < k_0^{II}.$$

Appendix C

Appendix for Chapter 3

Proof of Proposition 1. By (3.5), we have

$$w_1 = \frac{q_1 U^*}{(1+\rho)(1-e^{-q_1})} \text{ for } q_1 > 0,$$

and w_1 is not well-defined when $q_1 = 0$. So there is a one-to-one and negative relation between w_1 and q_1 when $q_1 > 0$. The maximization problem (3.7) is therefore equivalent to the following,

$$\pi_1^* = \max_{q_1 > 0} \{\rho \Phi_1(q_1) - q_1 U^*\} \tag{C.1}$$

Similarly, by (3.6), we have

$$w_2 = U^* \left[e^{-q_2} (F_{21}(\rho) + 1) + \frac{1}{q_2} \left(1 - e^{-q_2} - q_2 e^{-q_2}\right)(F_{22}(\rho) + 2) \right]^{-1} \text{ for } q_2 > 0$$

So the problem of (3.8) can also be re-written so that q_2 is the control variable,

$$\pi_2^* = \max_{q_2>0} \left\{ \Phi_{21}(q_2) F_{21}(\rho) + \Phi_{22}(q_2) F_{22}(\rho) - q_2 U^* - C \right\}. \tag{C.2}$$

The first-order conditions to (C.1) and (C.2) are

$$U^* \geq \rho e^{-q_1}, \tag{C.3}$$

$$U^* \geq e^{-q_2} F_{21}(\rho) + q_2 e^{-q_2} \left(F_{22}(\rho) - F_{21}(\rho) \right), \tag{C.4}$$

where the equalities hold when $q_1, q_2 > 0$. We focus on the situation where both small and large firms coexist, so we combine (C.3) and (C.4) at equalities and obtain the necessary condition for interior solutions (q_1^*, q_2^*),

$$q_1^* = q_2^* - \ln \left(\frac{1}{\rho} \left[F_{21}(\rho) + q_2^* \left(F_{22}(\rho) - F_{21}(\rho) \right) \right] \right), \text{ and } q_1^* > 0. \tag{C.5}$$

This also implies that $q_2^* > q_1^*$. Moreover, the necessary condition for coexistence requires $\pi^* = \pi_1^* = \pi_2^*$, which implies

$$\begin{aligned} & \rho \left(1 - e^{-q_1^*} - q_1^* e^{-q_1^*} \right) \\ = & \left(1 - e^{-q_2^*} - q_2^* e^{-q_2^*} \right) F_{22}(\rho) - \left(q_2^* \right)^2 e^{-q_2^*} \left(F_{22}(\rho) - F_{21}(\rho) \right) - C. \end{aligned} \tag{C.6}$$

These two equations give the unique solution (q_1^*, q_2^*) when it exists. Then (w_1^*, w_2^*) can be expressed as functions of (q_1^*, q_2^*) by using (3.5), (3.6), (C.3) and (C.4). *Q.E.D.*

Bibliography

[1] Acemoglu, D., U. Akcigit, N. Bloom, and W. Kerr (2011): "Innovation, Reallocation and Growth," *Working Paper*, University of Pennsylvania.

[2] Acemoglu, D. V., Carvalho, A., Ozdaglar, and A. Tahbaz-Salehim (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977-2016.

[3] Angeletos, G., and J. La'O (2013): "Sentiments," *Econometrica*, 81, 739–779.

[4] Arellano, C., Y. Bai, and P. Kehoe (2009): "Financial Markets and Fluctuations in Uncertainty," *Working Paper*, Federal Reserve Bank of Minneapolis.

[5] Atalay, E. (2013): "How Important Are Sectoral Shocks?" *Working Paper*.

[6] Bachmann, R., and C. Bayer (2011): "Uncertain Business Cycles - Really?" *NBER Working Paper*.

[7] Bachmann, R., and S. Elstner and E. Sims (2011): "Uncertainty and Economic Activity: Evidence from Business Survey Data," *Working Paper*, University of Michigan.

[8] Bachmann, R., R. Caballero, and E. Engel (2011): "Aggregate Implications

of Lumpy Investment: New Evidence and a DSGE Model," *Working Paper*, University of Michigan.

[9] Bachmann, R., and G. Moscarini (2011): "Business Cycles and Endogenous Uncertainty," *Working Paper*, University of Michigan.

[10] Balasubramanian, N., and J. Sivadasan (2009): "Capital Resalability, Productivity Dispersion, and Market Structure," *Review of Economics and Statistics*, Vol. 91, 547–557.

[11] Berger, D., and J. Vavra (2011): "Dynamics of the U.S. Price Distribution," *Working Paper*, Yale University.

[12] Bigio, S., and J. La'O (2013): "Financial Frictions in Production Networks," *Working Paper*.

[13] Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, Vol. 77, 623-685.

[14] Bloom, N., M. Floetotto, and N. Jaimovich (2010): "Really Uncertain Business Cycles," *Working Paper*, Stanford University.

[15] Brown, C. and J. Medoff, The Employer Size-Wage Effect, J. Polit. Economy 97 (1989) 1027–1059

[16] Burdett, K., S. Shi, and R. Wright (2001): "Pricing and Matching with Frictions," *J. Polit. Economy* 109, 1260-1085.

[17] Campbell, J. R., (1998): "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, Vol. 1, 371-408.

[18] Chugh, S. (2010): "Firm Risk and Leverage-Based Business Cycles," *Working Paper*, Univeristy of Maryland.

[19] Clementi, G.L., and D. Palazzo (2010): "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *Working Paper*, New York University.

[20] Cooper, R., and J. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, Vol. 73, 611-633.

[21] Conley, T., and B. Dupor (2003): "A Spatial Analysis of Sectoral Complementarity," *Journal of Political Economy*, 111, 311–351.

[22] Cohen, L., and A. Frazzini (2008): "Economic Links and Predictable Returns," *Journal of Finance*, 63, 1977-2011.

[23] Dhawan, R. (2001): "Firm Size and Productivity Differential: Theory and Evidence from a Panel of US Firms," *Journal of Economic Behavior & Organization*, Vol. 44, 269-293.

[24] Dunne, T., M. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, Vol. 19, 495-515.

[25] Davis, S., R. Faberman, and J. Haltiwanger (2010): "The Establishment-Level Behavior of Vacancies and Hiring," NBER working paper No. 16265 .

[26] Dupor, B., (1999): "Aggregation and Irrelevance in Multi-Sector Models," *Journal of Monetary Economics*, 43, 391–409.

[27] Eisfeldt, A.L., and A.A. Rampini (2006): "Capital Reallocation and Liquidity," *Journal of Monetary Economics*, Vol. 53, 369-399.

[28] Foerster, A., P. Sarte, and M. Watson (2011): "Sectoral versus Aggregate Shocks: A Structural Factor Anslysis of Industrial Production," *Journal of Political Economy*, 119, 1–38.

[29] Gabaix, X. (2011): The Granular Origins of Aggregate Fluctuations, *Econometrica*, 79, 733–772.

[30] Galenianos, M. (2012): "Learning About Match Quality and the Use of Referrals", Working Paper, Penn State University.

[31] Gertler, M., and S. Gilchrist (1994): "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics*, Vol. 109, 309-340.

[32] Hamilton, B.H. (2000): "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-employment." *Journal of Political Econnomy*, Vol. 108, 604–631.

[33] Hawkins, W., Competitive Search, Efficiency, and Multi-worker Firms, Working Paper, University of Rochester, 2011.

[34] Herranz, N., S. Krasa, and A. Villamil (2009): "Small Firms in the SSBF," *Annals of Finance*, Vol. 5, 341-359.

[35] Higson, C., S. Holly and P. Kattuman (2002): "The Cross-sectional Dynamics of the US Business Cycle: 1950–1999," *Journal of Economic Dynamics & Control*, Vol. 26, 1539-1555.

[36] Higson, C., S. Holly, P. Kattuman, and S. Platis (2004): "The Business Cycle, Macroeconomic Shocks and the Cross-section: the Growth of UK Quoted Companies," *Economica*, Vol. 71, 299-318.

[37] Hopenhayn, H. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, Vol. 60, 1127-1150.

[38] Kaas, L., and P. Kircher, Efficient Firm Dynamics in a Frictional Labor Market, Working Paper, LSE, 2011.

[39] Lester, B.(2010) "Directed Search with Multi-vacancy Firms," *J. Econ. Theory* 149, 2108-2132.

[40] Long, J. B., and C. I. Plosser (1983): "Real Business Cycles," *Journal of Political Economy*, 91, 39–69.

[41] Lorenzoni, G. (2011): "News and Aggregate Demand Shocks," *Annual Review of Economics*, 3, 537-557.

[42] Horvath, M. (1998): "Cyclicality and Sectorial Linkages: Aggregate Fluctuations From Sectorial Shocks," *Review of Economic Dynamics*, 1, 781–808.

[43] Horvath, M. (2000): "Sectorial Shocks and Aggregate Fluctuations," *Journal of Monetary Economics*, 45, 69–106.

[44] Jackson, M. O. (2008): *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

[45] Jaimovich, N., and M. Floetotto (2008): "Firm Dynamics, Markup Variations, and the Business Cycle," *Journal of Monetary Economics*, Vol. 55, 1238-1252.

[46] Jovanovic, B. (1987): "Micro Shocks and Aggregate Risk," *Quarterly Journal of Economics*, 102, 395–409.

[47] Kehrig, M. (2011): "The Cyclicality of Productivity Dispersion," *Working Paper*, Northwestern University.

[48] Kelly, B, Lustig, H, and Van Nieuwerburgh, S. (2013): "Firm Volatility in Granular Networks," *Working Paper*.

[49] Khan, A., and J. Thomas (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, Vol. 76, 395–436.

[50] Lee, Y., and T. Mukoyama (2008): "Entry, Exit and Plant-level Dynamics over the Business Cycle," *Working Paper*, Federal Reserve Bank of Cleveland.

[51] Meyer, M. (1994): "The Dynamics of Learning with Team Production: Implications for Task Assignment," *Quart. J. Econ.* 109, 1157-1184.

[52] Montgomery, J. (1991): "Equilibrium Wage Dispersion and Interindustry Wage Differentials", *Quart. J. Econ.* 106, 163-179.

[53] Moskowitz, T.J., and A. Vissing-Jorgensen (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *American Economic Review*, Vol. 92, 745–778.

[54] Oberfield, E. (2013), "Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture," Working Paper.

[55] Oi, W., and T. Idson (1999): “Firm Size and Wages”, in: O.C. Ashenfelter and D. Card, (Eds.), Handbook of Labor Economics, Elsevier Amsterdam, pp. 2165–2214.

[56] Ortega, J. (2001): “Job Rotation as a Learning Mechanism”, *Management Science* 47, 1361-1370.

[57] Ouyang, M. (2009): "The scarring effect of recessions," *Journal of Monetary Economics*, Vol. 56, 184-199.

[58] Papageorgiou, T. (2011): Large Firms and Internal Labor Market, Working Paper, Penn State University.

[59] Peters, M. (1991) “Ex Ante Price Offers in Matching Games Non-Steady States,” *Econometrica* 59, 1425-1454.

[60] Quadrini, V. (2009): "Entrepreneurship in Macroeconomics," *Annals of Finance*, Vol. 5, 295-311.

[61] Shea, J. 2002: "Complementarities and Co-movements," *Journal of Money, Credit, and Banking*, 34, 412–433.

[62] Shi, S. (2002): “Product Market and the Size-wage Differential,” *Int. Econ. Rev.* 43, 21-45.

[63] Tan, S. (2012): “Directed Search and Firm Size,” *Int. Econ. Rev.* 53, 95-113.

[64] Vereshchagina, G., and H. Hopenhayn (2009): "Risk Taking by Entrepreneurs," *American Economic Review*, Vol. 99, 1808-1830.