Short Lived Shocks and Aggregate Productivity

Ezra Oberfield

University of Chicago

PRELIMINARY AND INCOMPLETE

January 11, 2010

When investment is partially irreversible, a short lived spike in the real interest rate can cause a decline in measured total factor productivity. The increased interest rate widens the inaction region, causing firms to be less responsive to productivity shocks and leading to a less efficient allocation of capital. In 1982, the Chilean manufacturing sector suffered a large contraction following a spike in real interest rates and a large run up in real wages. I test the model using microdata on Chilean manufacturing firms with mixed success. The timing of changes in the dispersion of average product of capital is consistent with the model. I find that this increased dispersion has a substantial impact on total factor productivity, but is unlikely to be the main determinant of changes in TFP during the recession.

In 1982, Chile suffered a classic sudden stop. Output declined, unemployment rose, and capital inflows stopped. The contraction of the manufacturing sector was particularly severe, with output shrinking more than twenty percent and measured total factor productivity falling by more than ten percent. Among the many factors that played a role, I explore the impact of two large shocks that hit the Chilean economy: a spike in the real interest rate and a run-up in real wages. The real interest rate rose substantially in 1981 and 1982, partially due to the increase in the US fed funds rate and partially due to an increased default risk. By law, nominal wages were indexed to past changes in the CPI, so a declining rate of inflation led to an increase in real wages.

When these shocks are expected to be short-lived, they can exacerbate investment and hiring irreversibilities, causing firms to be less responsive to productivity shocks and slowing the reallocation of resources from relatively low to high productivity firms. The cross sectional dispersion of marginal product slowly increases, leading to a less efficient allocation of capital and a decline in measured TFP. I assess the model using microdata on Chilean firms to investigate two questions: (i) Did these shocks have an impact on TFP? (ii) If so, can these shocks account for the large movements in TFP during the recession?

I find that increased dispersion in average products across firms led to a substantial decline in measured TFP in 1981 and 1982, and the timing is consistent with prediction of the model. However, it seems unlikely that changes in dispersion were the main contributor to changes in TFP, though this result is sensitive to the exact empirical specification.

A permanent increase in the real interest rate reduces option value of waiting to invest or disinvest, as the firm puts less weight on future cashflows. In contrast, a spike that is expected to be short-lived widens the region of inaction. Firms that were on the verge of investing now put this off until the real interest rate falls. Firms are also reluctant to disinvest, an action they would regret once the real interest rate returns to normal. As a consequence, the (s,S) bands widen, slowing reallocation and lowering aggregate productivity. Measured TFP falls as the share of capital and output accounted for by relatively low productivity firms rises.

The increased dispersion of marginal products implied by the widening inaction region may understate the short run decline in aggregate productivity. Specifically, if firms invest much more frequently than they disinvest, aggregate productivity overshoots. Initially the distribution of firms is concentrated in the state space close to the investment threshold. At the onset of a shock, this

threshold shifts up. The distribution of firms fans out across the inaction region before becoming relatively concentrated near the new investment threshold. Consequently, there is more dispersion in marginal products in the short run than in the long run, leading to an overshoot of aggregate productivity.

The effect of a temporary increase in the real wage depends on how flexibly firms respond to the shock. If there are no hiring frictions, the wage increase will widen the investment inaction region, as the wage increase has the same effect as a temporary productivity shock. To the extent that hiring frictions limit the adjustment of employment, inaction region for labor adjustment will widen inducing dispersion in the marginal product of labor. In either case, the wage shock lowers aggregate productivity.

A key prediction of the model is that the distribution of marginal product of capital becomes more dispersed after the onset of an interest rate shock. While marginal products cannot generally be directly measured, by assuming a Cobb Douglas production function I can measure the average product which is proportional to the marginal product. I use plant level data from the annual Chilean Manufacturing Survey to analyze changes in the dispersion of average products of capital and labor over time. Using two different measures, the dispersion in the average products of capital and labor rise following the shocks.

Another prediction is that the cross-sectional correlation of revenue and capital falls after an interest rate shock. In Chile there is a clear decline in this correlation in 1982, suggesting a less efficient allocation of resources. In addition, the correlation between changes in revenue with changes in capital also declines in 1982, although it remains higher in the early 1980s than in later years. This is evidence that reallocation of resources across firms slowed in 1982.

These findings are consistent with the hypothesis that irreversibilities played an important role in determining changes in the allocation of resources across firms. To quantify the impact of these changes on measured TFP, I develop a decomposition of aggregate TFP using wedge accounting in the spirit of Hsieh and Klenow (2009) and Chari et al. (2007). This measures the wedge between the actual marginal product of a firm and the theoretical marginal product in a world with no optimization frictions. I separate TFP into two components, one that depends only on firm level productivity parameters, and one that summarizes the contribution of the wedges within

and between sectors[1]. The model predicts that the after the onset of the interest rate shocks, the second term (which depends on the cross-sectional distribution of wedges) would gradually decline, and then rise once the interest rate returns to a low level.

For my preferred specification, I find that a worsening allocation of resources decreased TFP by 8% over 1981 and 1982 and raised TFP sharply in 1983. The timing of these movements supports the hypothesis that the spike in real interest rates gradually raised dispersion and that this dispersion led to a large decline in TFP. However, the timing also suggests that the increased dispersion was not the main driver of movements in TFP during the recession.

The decomposition presented here has several appealing properties. First, the decomposition is done in a way that is compatible with business cycle accounting at the aggregate level and can be used to complement the type of analysis performed by Chari et al. (2005).

Second, only the nominal values of revenue and payment to the various factors of production are needed[2]. In a high inflation environment, deflators may be poorly measured, and can introduce additional noise into measurement. Further, using expenditures on inputs helps to control for the quality those inputs. For example, the wage bill may be a better measure of the efficiency units of labor than the number of workers if the workers are heterogeneous.

Third, noisy data is less likely to be a problem for time series comparisons than for cross-country comparisons. A critique of the Hsieh and Klenow (2009) methodology is that the increased measured dispersion in average products in developing countries may simply be an artifact of noisier data in those countries, as data collection may be less reliable. It is more plausible that the reliability of data collection is more or less consistent across years within the same country than it is across countries.

Several studies have argued that plant level irreversibilities are not relevant for understanding aggregate fluctuations in general equilibrium because prices would adjust and mute any additional effect[3]. For the purposes of this paper, I focus less on the determination of these prices than

[1] The decomposition has nothing to say about the source of these wedges; they could stem from adjustment costs, financing frictions, information issues, measurement error, or a host of other factors. Nevertheless, the time series of the distribution of the wedges can give clues about the source of aggregate fluctuations.

[2] Revenue, payments to labor, and expenditures on intermediate inputs are directly reported in the census. One must impute payments to capital.

[3] The most closely related paper, Veracierto (2002), argues that plant level investment irreversibilities do not play a significant role in standard real business cycle models. Thomas (2002) and Khan and Thomas (2003) report that consumption smoothing essentially eliminates the impact of lumpy adjustment costs on aggregate fluctuations. House (2008) explains this most clearly: Firms are very willing to adjust the timing of investment in response to

on the decisions of individual firms that take the evolution of these prices as given. In addition, the economic forces that drive these irrelevance results are less likely to be important during the Chilean recession. To the extent that a small open economy like Chile faces a more elastic supply of capital than a closed economy, irreversibilities will have a larger impact on aggregate outcomes. Also, real wages were indexed to inflation and hence could not adjust to the increasing supply of available labor.

The model has a similar flavor as the series of papers about uncertainty shocks, including Bloom (2007), Bloom et al. (2007), and Bloom et al. (2009). Those papers model changes in the variance of firm level shocks and argue that a doubling of the magnitude of firm level shocks can lead to business cycle size drops in output and TFP. In that framework, the shocks cause firms to be more cautious in investment decisions as the option value of waiting increases. This causes a widening of (s,S) bands and a slowing of reallocation from high to low productivity firms. Bloom et al. (2009) develop an uncertainty index and argue that shocks to this index lead to drops in output.

I show an additional causal link between changes in uncertainty and a widening of the (s,S) bands. An increase in aggregate uncertainty can raise the risk premium and hence real interest rates. If the uncertainty index is correlated with aggregate uncertainty there is problem identifying the causal channel. It is likely that both forces are at play, and I view these channels as complementary.

One way to distinguish between changes in firm level uncertainty and changes in the interest rate induced by aggregate uncertainty is to study plant level exit behavior. An increase in firm level uncertainty increases the option value of waiting for news, making firms less likely to exit. An increase in the interest rate reduces that option value, increasing exits.

Eisfeldt and Rampini (2006) show that capital reallocation in the US is procyclical and argue that this implies a large countercyclical reallocation cost. The model presented here gives an alternative explanation for the temporal pattern of reallocation: that shocks are expected to be short-lived.

small changes in price (hence the large inaction region), making demand for investment very elastic. If the elasticity of supply is less elastic (e.g., due to desire for smooth consumption) the quantity of investment will be pinned down by the supply curve. This makes the distribution of firms in the inaction region less relevant for aggregate behavior. Lastly, Veracierto (2008) argues that hiring and firing frictions reduce the effect of aggregate fluctuations on output, productivity, and employment.

Several studies have sought a connection between rising interest rates and declines in productivity in emerging markets. Neumeyer and Perri (2005) propose a model in which there is a lag between production of goods and revenue, so that firms must borrow to pay workers. In that setting a higher interest rate raises the effective cost of labor. Meza and Quintin (2007) conjecture that the drop in measured TFP during Mexico's financial crisis can be accounted for by a decrease in utilization and decrease in labor effort. These models differ from the one presented here in that they predict declining measured productivity for each individual firm.

A growing literature explores the link between the allocation of resources across plants and measured TFP. Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Banerjee and Duflo (2005) propose plant-level misallocation as an explanation for cross country income differences. This literature has mainly focused on cross-country differences. This paper extends this empirical work by studying fluctuations in output at business cycle frequencies.

Methodologically this paper builds on a large literature concerned with investment under uncertainty. Most specifically Abel and Eberly (1996) characterized the optimal policy of a firm that faced investment with costly reversibility. Oberfield (2008) extended this work by adding endogenous entry and exit. This is non-trivial extension because the solution method used in the literature relies on the homogeneity of the value function, which is not compatible with a model in which low productivity firms exit. Guo et al. (2005) study the investment policy of a firm facing complete irreversibility and shifts between regimes that differ in the drift and variance of productivity shocks. The model also shares features with Campbell and Fisher (2000) who study the decisions of a firm subject to shifts between high and low wage regimes, though they were primarily interested in the cyclicality of job creation and destruction.

In Section 1, I discusses the economic environment and the impact of the recession. In section 2, I describe the impact of short lived shocks in a model with irreversibilities. I describe the data and present some preliminary evidence in section 3. Lastly I develop the TFP decomposition and apply it to the Chilean data in section 4.

1 Reform and Recession in Chile

In the 1970's and early 1980's, Chile's economic landscape changed drastically. Pinochet came to power in 1973 and, along with the Chicago Boys, instituted a number of economic reforms. In 1974 the Chilean government privatized the nation's banks. Several were subsequently bailed out after poor performance in 1976-77, but for the most part the banks were in private hands until the crisis in the early 1980's, when several banks had to be bailed out again.

To tackle runaway inflation, the government instituted a series of planned currency devaluations relative to the dollar in 1978 that culminated with a fixed exchange nominal exchange rate with the dollar in 1979. The fixed rate remained in place until June 1982, in the middle of the recession.

From 1973 through 1979, nominal wages for some workers were periodically adjusted for increases in the CPI through legislation. In 1979, this was extended to all workers. Any new wage offers had a floor of the previous wage multiplied by the change in CPI since the previous wage offer. Another labor reform in 1978 attempted to reduce employment frictions by eliminating the "just cause" doctrine and capping severance payments at five months. Hiring frictions remained sizable, as severance was equal to one month for every year worked (up to this five month cap).

In 1973, tariffs were large and varied widely, averaging 105%. There were also many non-tariff trade barriers. From 1973-1979, these trade barriers were gradually reduced to a uniform tariff of 10%. Some argue that this led to a high exit rate among small plants (see Pavcnik (2002)).

1.1 Recession

The Chilean economy suffered a large contraction in 1982. Figure 1(a) shows the large decline in value added and TFP for the Chilean economy as a whole with data from the World Bank. As can be seen in figure 1(b), the contraction of the manufacturing sector was particularly severe.

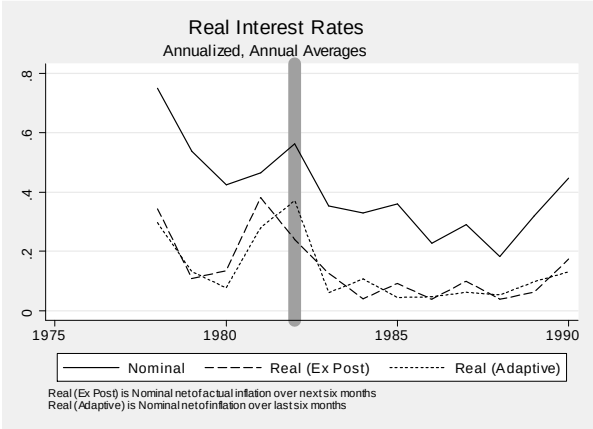
Inflation, which was in the triple digits during much the 1970's, began to decline in the late 1970's and early 80's. As can be seen in figure 2, real interest rates spiked in 1981 and remained high in 1982. While I have no direct measure of the real interest rate, I can subtract expected inflation from the nominal rate. I compute two estimates of expected inflation estimates, one that uses the actual ex post inflation during the period, and one that uses the lagged inflation. Actual expectations of inflation were most likely in between these numbers, and for the remainder of the

Figure 1: Value Added and Total Factor Productivity

(a) Chilean Economy (b) Manufacturing Sector

paper I use an average of the two.

Figure 2: Real Interest Rates



Real wages also spiked during the recession. Because wages were indexed to previous inflation, the declining inflation led to an increasing real wage rate even when labor productivity was stagnant. Figure 3 shows the large rise in real wages, which was especially pronounced in the manufacturing sector.

2 Irreversibilities and Short-Lived Shocks

A short-lived shock to the real interest rate impacts the investment decisions of individual plants, and these micro decisions can aggregate to lower measured TFP. I first analyze the decisions of a

Figure 3: Real Wages

single price taking firm subject to idiosyncratic productivity shocks, partial investment irreversibility, and a stochastic real interest rate. I then compute the impulse response for a sector composed of a large number of firms and show how the individual decisions aggregate up to affect output and aggregate productivity.

Depending on how flexibly firms adjust labor supply, short lived shocks to the real wage can amplify the effect of a real interest rate shock. I briefly discuss this and show how the aggregate response to the shock may differ from a real interest shock.

In the section below I describe the general model in which there is a fixed production cost that induces low productivity firms to exit. When I discuss an environment with a constant interest rate, it is useful to see how the exit threshold depends on the interest rate. However, when I study an environment with a stochastic interest rate, I set the fixed cost to zero, as the assumptions that allow me to solve the model with endogenous exit become too costly.

2.1 Real Interest Rate Shocks

A firm has access to a technology to produce a consumption good using capital and labor. The technology undergoes idiosyncratic productivity shocks. Capital purchases are only partially reversible; the firm can purchase capital at price p^+ and resell it at p^-. For concreteness, I assume this wedge between purchase and resale price arises from the specificity of the capital. The firm must pay a fixed production cost anytime it is actively producing. At any point in time, the firm can choose to exit, shutting down production and selling off all remaining capital. In addition, firms die at an exogenous rate χ.

A plant with productivity $\tilde{A}$, capital K, and labor L produces $\tilde{A}K^{\alpha}L^{1-\alpha}$. The firm faces and isoelastic demand curve, $Y = DP^{-\sigma}$. Flow profits can then be written as

$$\pi = P\tilde{A}K^{\alpha}L^{1-\alpha} - wL - f$$

where f is the fixed production cost. If the firm chooses P and L optimally, flow profits can be rewritten as

$$\pi\left(A, K\right) = \frac{1}{1-\theta}A^{\theta}K^{1-\theta} - f$$

where $A \propto \tilde{A}^{\sigma-1}$ and $\theta = \frac{1}{\alpha(\sigma-1)+1}$.

Productivity follows a geometric Brownian motion, $d\log A = \mu dt + \varepsilon d\omega$ where ω is a standard Weiner process. The real interest rate $r\left(t\right)$ follows a Markov process independent of ω.

A firm's policy is consists of three diffusions that are progressively measurable with respect to the realized shock process. The first and second are cumulative investment I_t^+ and disinvestment I_t^-. These processes are right continuous and weakly increasing, $dI_t^+, dI_t^- \geq 0$ but not differentiable. I use this notation to allow for the possibility of lumpy investment or disinvestment. Changes in capital stock can be written as $dK_t = -\delta K_t + dI_t^+ - dI_t^-$, so that the capital stock is $K_t = e^{-\delta t}K_0 + \int_0^t e^{-\delta(t-\tau)}\left(dI_\tau^+ - dI_\tau^-\right)$. The capital stock must be nonnegative. Third, the firm chooses a stopping time T at which it shuts down production and exits. Let $\mathcal{P}$ be a policy (I^+, I^-, T).

The firm chooses an investment policy to maximize the present discounted value of profits net of investment costs. For a given policy $\mathcal{P}$, the present value to a firm of following that policy is $W\left(A, K, r; \mathcal{P}\right)$. W satisfies

$$W\left(A_0, K_0, r_0; \mathcal{P}\right) = \mathbb{E}\left[\int_0^T e^{-\int_0^t r(\tau)d\tau - \chi t}\left[\pi\left(A_t, K_t\right)dt - p^+ dI_t^+ - p^- dI_t^-\right]\right]$$

The value of the optimal policy can be written as

$$V\left(A_0, K_0, r_0\right) = \max_{\mathcal{P}} W\left(A_0, K_0, r_0; \mathcal{P}\right)$$

The wedge between the purchase and resale prices creates a kink in the cost schedule of net investment at zero. For a non-trivial range of productivity, the firm will maximize profit by

choosing zero net investment. This range is a region of inaction, and the optimal policy will take the form of threshold rules. For any K and r, there are thresholds $\bar{A}(K,r)$ and $\underline{A}(K,r)$ such that if a firm has productivity above $\bar{A}(K,r)$ it invests while if productivity drops below $\underline{A}(K,r)$ the sells capital.

2.2 Constant Interest Rate

Here I consider the firm's problem when the interest rate process is degenerate and only takes a single value. For this subsection I also abstract from depreciation, setting $\delta = 0$. In this special case, the optimal policy can be characterized by hand as the model can be solved almost completely in closed form. This allows for a clear characterization of the solution and comparative statics, and helps give intuition for the various channels through which the interest rate affects the actions of the firm.

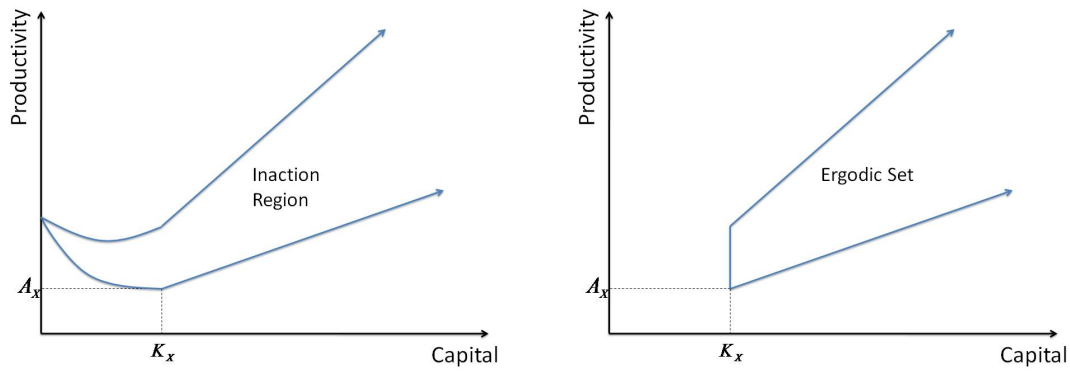
Figure 4 shows the optimal policy and the resulting ergodic set. At a given point in time a firm is located at a single point on this graph. Productivity shocks shift the firm up or down. When a firm in the inaction region reaches the upper threshold, it purchase capital, which shifts the firm to the right, and a firm that sells capital at the lower threshold moves to the left. Outside the region of inaction, a firm immediately invests or disinvests enough capital to return to the region of inaction. In addition, when productivity drops low enough, the firm exits and sells all remaining capital. The exit point is labeled (K_X, A_X). The section of the inaction region to the right of the exit point forms an ergodic set. A firm with capital greater than K_X that receives a series of negative productivity shocks will reduce its capital holdings until K_X; further adverse productivity shocks will cause the firm to exit.

2.2.1 Optimal Policies

For any policy with threshold rules of this type, the value function satisfies the Hamilton-Bellman-Jacobi equation when the firm is in the region of inaction:

$$rW(A,K) = \frac{1}{1-\theta}A^{\theta}K^{1-\theta} - f + \frac{1}{dt}\mathbb{E}\left[dW(A,K)\right] \tag{1}$$

Figure 4:



(a) Optimal Policy (b) Ergodic Set

where, by Ito's Lemma $\frac{1}{dt}\mathbb{E}\left[dW\right] = \mu A W_A + \frac{\varepsilon^2}{2} A^2 W_{AA}$. For each K this is an ODE which has the general solution

$$W\left(A,K\right) = \frac{1}{1-\theta}\frac{1}{H\left(\theta\right)}A^{\theta}K^{1-\theta} - \frac{f}{r} + \sum_{i=1,2}\tilde{c}_i\left(K\right)A^{s_i}$$

where $H\left(s\right) = r - \mu s - \frac{\varepsilon^2}{2}s\left(s-1\right)$ and s_1 and s_2 are the roots of $H\left(s\right) = 0$. Note that the constants of integration $\tilde{c}_1\left(K\right)$ and $\tilde{c}_2\left(K\right)$ are functions of K, not simply numbers.

For any investment and disinvestment thresholds, $\bar{A}\left(K\right)$ and $\underline{A}\left(K\right)$, the value function will satisfy the following smooth pasting conditions:

$$\begin{aligned} p^{+} &= W_K\left(\bar{A}\left(K\right),K\right) = \frac{1}{H\left(\theta\right)}\bar{A}\left(K\right)^{\theta}K^{-\theta} + \sum_{i=1,2}\tilde{c}_i'\left(K\right)\bar{A}\left(K\right)^{s_i} \\ p^{-} &= W_K\left(\underline{A}\left(K\right),K\right) = \frac{1}{H\left(\theta\right)}\underline{A}\left(K\right)^{\theta}K^{-\theta} + \sum_{i=1,2}\tilde{c}_i'\left(K\right)\underline{A}\left(K\right)^{s_i} \end{aligned}$$

The *optimal* thresholds also satisfy the following supercontact conditions.

$$\begin{aligned} 0 &= W_{KA}\left(\bar{A}\left(K\right),K\right) = \theta\frac{1}{H\left(\theta\right)}\bar{A}\left(K\right)^{\theta-1}K^{-\theta} + \sum_{i=1,2}s_i\tilde{c}_i'\left(K\right)\bar{A}\left(K\right)^{s_i-1} \\ 0 &= W_{KA}\left(\underline{A}\left(K\right),K\right) = \theta\frac{1}{H\left(\theta\right)}\underline{A}\left(K\right)^{\theta-1}K^{-\theta} + \sum_{i=1,2}s_i\tilde{c}_i'\left(K\right)\underline{A}\left(K\right)^{s_i-1} \end{aligned}$$

Appendix A.1 gives a heuristic derivation of these conditions. The following proposition shows

that the optimal thresholds are linear (see appendix for proof).

Proposition 1 *For $K > K_X$, the upper and lower thresholds take the form $\bar{A}(K) = \bar{Z}K$ and $\underline{A}(K) = \underline{Z}K$*

The linear threshold rules are essentially the same as those derived by Abel and Eberly (1996). This may appear surprising because it implies that neither the fixed cost of production nor the exit decision to have an impact on the thresholds. One might think that changes in the value function near the exit point might percolate up to effect the investment and disinvestment decisions at different levels of capital.

The thresholds depend only on the marginal value of capital. When a firms invests, it receives a unit of capital and the option to disinvest. The marginal value of the capital is the expected change in cash flows before that option to disinvest is exercised. Since there is no depreciation, the firm is certain to exercise the option to disinvest before it makes a decision to exit. Therefore the marginal value of the unit of capital is independent of the parameters that govern the exit decision[4].

We now turn to the determinants of the size of the region of inaction. Let the wedge between the purchase and resale price of capital be $B \equiv \frac{p^+}{p^-}$. Also let $X \equiv \bar{Z}/\underline{Z}$ be a measure of the size of the inaction region (which will later be an important determinant of the dispersion of capital holdings across firms). As shown in appendix A.2, X satisfies the equation

$$G(X, B, s_1) = G(X, B, s_2) \tag{2}$$

where $G(X, B, s) \equiv \frac{s-\theta}{s} \frac{X^s - B}{X^s - X^\theta}$, and the thresholds satisfy

$$\bar{Z}^\theta = rp^+ G\left(X^{-1}, B^{-1}, s\right) \tag{3}$$

$$\underline{Z}^\theta = rp^- G(X, B, s) \tag{4}$$

[4]Formally, the fixed cost and the exit decision affect the level of the value function but not its slope. The optimality conditions are homogeneous of degree zero in productivity and capital even though the value function is not HOD 1. Without depreciation, the HJB equation and boundary conditions can be solved separately for each K. Allowing for depreciation would add a drift term $-\delta K W_K$ to the HJB equation, so that this would become a PDE, and hence the boundary conditions for a given K would not be sufficient to solve for the thresholds at that K. This is natural as the state of the firm when the firm disinvests that particular unit of capital is no longer pinned down.

A firm will exit when the value of staying in business equals the value of shutting down. This is the value matching condition

$$W\left(A_X, K_X\right) - p^{-} K_X = 0$$

Where K_X and $A_X \equiv \underline{A}\left(K_X\right)$ are the level of productivity and capital at which the firm chooses to shut down. Appendix A.4 shows a heuristic derivation of a second condition that the optimal thresholds must satisfy:

$$W_A\left(A_X, K_X\right) = 0$$

With these two equations, we can solve for the two unknowns A_X and C_2. One can verify that

$$\begin{aligned} A_X &= \frac{1-s_1}{s_1} \frac{\underline{Z} f}{r p^{-} - \frac{\underline{Z}^{\theta}}{1-\theta}} \\ C_2 A_X^{s_2} &= \frac{1}{1-s_2} \frac{f}{r} \end{aligned}$$

As the fixed production cost increases, the exit threshold rises as firms are more willing to exit.

2.2.2 Comparative Statics

We now turn to the role of the interest rate in determining the thresholds. When the interest rate rises, the user cost of capital increases, making capital less attractive. This shifts both thresholds up. In addition, a higher interest rate reduces the option value of waiting to invest or disinvest. The firm places more weight on current cash flows, so changes in future productivity become less important. Through this channel, an increase in r has the same effect as proportional decreases in the magnitudes of the drift and variance of productivity shocks, μ and ε[5]. Because future cashflows are discounted at a higher rate, the future movements in the productivity are less important for current decisions, narrowing the region of inaction. See appendix A.3 for details.

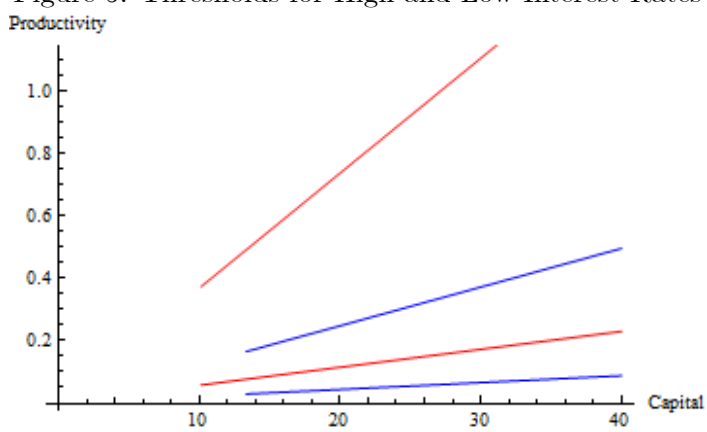
At a higher interest rate, the exit threshold is increased for a similar reason. A firm thinking about the possibility of exit may choose to stay in business because of the possibility of receiving a

[5] As can be seen from (2), (3), and (4), μ and ε^2 only enter the determination of size of the inaction region through roots of the characteristic polynomial, $H(s) = r - \mu s - \frac{\varepsilon^2}{2} s(s-1)$. By simply dividing $H(s)$ by r (leaving the roots unchanged) one can see that an increase in r has an identical effect on s_1 and s_2 as a decrease in the magnitudes of μ and σ^2.

series of positive productivity shocks and becoming profitable. However, when the interest rate is higher, the option value of waiting for these positive productivity shocks declines, making the firm more likely to exit.

Figure 5 shows the thresholds above and below ergodic set for two different interest rates. The blue lines are the thresholds for a firm facing a low interest rate while the red lines are those for a high interest rate. When the interest rate is higher, the size of the inaction region (ratio of $\bar{Z}/\underline{Z}$) is smaller and the exit threshold, A_X, is higher.

Figure 5: Thresholds for High and Low Interest Rates



2.3 Regime Shift

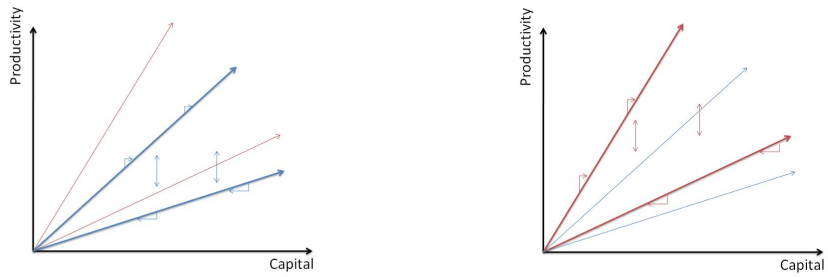
Now I allow for the possibility that the interest rate can change. I model real interest rate as a two state Markov process and assume the transition probability follows a Poisson law. The optimal policy will still be characterized by separate thresholds in each regime, and the thresholds will reflect the possibility of a change in interest rate.

Let r_i be the interest rate in state i with $r_1 < r_2$. Let λ_i be the Poisson arrival rate of a transition from state i to state $-i$, so the probability of a transition before time s is $e^{-\lambda_i s}$.

Figure 6 below shows the thresholds for the two regimes. While in one regime, the productivity shocks shift the firm up or down. When the firm's productivity reaches the upper threshold it invests incrementally, pushing the firm to the right. Similarly, disinvestment at the lower threshold pushes the firm to the left. When a regime shift happens, if the firm finds itself in between the thresholds of the new regime, it takes no action. If it finds itself above the investment threshold,

it immediately invests until it is back in the inaction region for the new regime[6].

Figure 6: Thresholds for the two Regimes



Low Interest Rate Regime High Interest Rate Regime

At this point I abstract away from endogenous exit, setting $f = 0$, and allowing for depreciation, $\delta > 0$. One can solve for the optimal thresholds by solving a system of nonlinear equations, as described in appendix B.

2.3.1 Duration of the Shock

An important feature of a real interest shock is its expected duration. Consider a firm that is at the investment threshold of the low interest regime when the shock hits. The increased user cost of capital makes investment less attractive. When the regime switches back to a low interest rate, the firm will want to invest, but while the interest rate is high it makes sense to wait.

Alternatively consider a firm at the disinvestment threshold when the high interest rate regime arrives. Again, the higher user cost of capital makes disinvest more attractive. However, since the shift is short lived, the firm must take into account that when the regime shifts back the capital will again be valuable. Thus while disinvestment is more attractive, it is only marginally more so because the high interest rate is expected to be temporary.

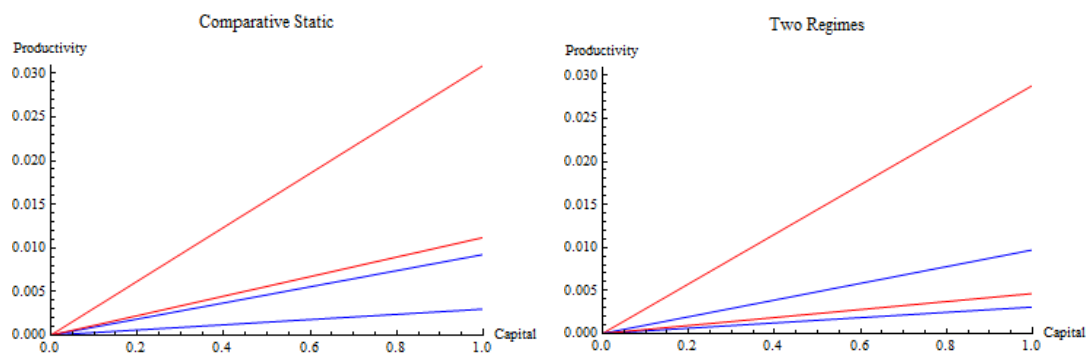
The difference comes from the asymmetry in investment technology: waiting to invest is easily reversible, while disinvestment is not. If a firm disinvests, it can only reinvest at a high cost.

[6] This lumpy investment behavior when an interest rate shock hits is an artifact of the modelling assumption of large jumps in the interest rate. A continuous interest rate process would lead to continuous changes in investment thresholds and hence more smooth investment patterns.

Because the high interest rate regime is short lived, the firm should only take actions that are easily reversible once the regime shifts back.

Figure 7 show the effect of short duration of the interest rate spike. These figures are calibrated so that the two interests rates are 10% and 20%. λ_1 is set .05, so that a sudden stop happens once every 20 years, whereas λ_2 is set to 1, so that it lasts one year on average. The left figure shows the thresholds when the interest regime is expected to last forever. In the figure to the left, firms account for the fact that the high interest regime will not last long, and the disinvestment margin practically coincides with that of the low interest regime.

Figure 7:



(a) $\lambda_1 = 0$, $\lambda_2 = 0$ (b) $\lambda_1 = .05$, $\lambda_2 = 1$

This mechanism can be illustrated most clearly in the extreme case of $\lambda_2 \to \infty$. This means that when the high interest rate shock occurs, it lasts for an instant and the economy immediately switches back to a low interest rate. This limiting case is roughly identical to a world with only a low interest rate, but it is useful to consider the behavior of the thresholds in each regime.

In the low interest rate regime, the thresholds in the limiting economy coincide with those of the world with only the low interest rate. In the high interest rate regime, the lower threshold also coincides with the disinvestment threshold in the low interest rate world. However, the investment threshold remains higher than that of the low interest rate regime. Firms are more reluctant to invest in the high interest rate regime, even though that regime is fleeting.

More formally one can take the following limit:

$$\lim_{\lambda_2 \to \infty} (\underline{Z}_2 - \underline{Z}_1) = 0 \quad \text{and} \quad \lim_{\lambda_2 \to \infty} \left(\bar{Z}_2 - \bar{Z}_1\right) > 0$$

See appendix B.1 for details. The limit of the marginal value of capital in the short lived, high interest rate regime is subtle. At the upper threshold for regime one, $\bar{Z}_1$, the marginal value of capital in regime 2 converges to p^+, $\lim_{\lambda_2 \to \infty} v_2\left(\bar{Z}_1\right) = p^+$, which seemingly implies that the firm should invest. However, this hides the fact that the high interest rate is extremely short-lived, so the investment decision while the interest rate is still high has an infinitesimal effect on the marginal value of capital. If we scale by the duration of the regime, we get the following limit.

$$\lim_{\lambda_2 \to \infty} (r_2 + \lambda_2)\left(v_2\left(\bar{Z}_1\right) - p^+\right) = (r_1 - r_2)\, p^+ < 0$$

It follows that the marginal value of capital is not high enough to induce the firm to invest in this regime. Hence the investment threshold must be higher than $\bar{Z}_1$.

This asymmetric reaction to short lived shocks in the face of irreversibilities is robust to many different kinds of shocks.

2.3.2 A Dynamic Sector

Consider a mass of firms producing imperfect substitutes with elasticity of substitution ε. These firms are a negligible part of a large sector, and take sector demand as given. Formally, firms face demand curves of $Y_i = DP_i^{-\varepsilon}$, and changes in the prices charged by these firms have a negligible impact on D. This assumption might be applicable to a sector in a small open economy that is a negligible part of a world industry and is unable to affect the world price of the industry aggregate[7][8].

The firms have identical production functions but differ in productivity. A flow of new firms, E_t, enters at productivity A_0 each instant. These firms invest enough to jump to the investment threshold.

At any point in time, one can compute the Solow residual for a representative firm for this sector.

[7] In a perfect competition setting, this would be equivalent to assuming that there is a world price for the good.

[8] An alternative assumption would be that D is an increasing function of the price index for the collection of firms, P. In this case the change in D would depend on the change in TFP and on the elasticity of demand for the sector aggregate. A fall in TFP for the collection of firms would raise the price index, shifting out the demand curve for individual firms. This would raise the productivity of capital, dampening the effect of the increased user cost of capital.

In response to an interest rate spike, TFP falls only gradually. The dampening would be small on impact but would grow over time.

Real output for the sector, Y_s would be computed as $Y_s = \frac{\int P_i Y_i}{P_s}$, where $P_s = \left(\int P_i^{1-\varepsilon}\right)^{1/(1-\varepsilon)}$ is an ideal price index for the sector. If K_s and L_s represent total capital and labor used in the sector, the Solow residual would be

$$TFP = \frac{Y_s}{K_s^{\alpha} L_s^{1-\alpha}}$$

where α is the capital share taken from the national accounts.

The Solow residual can be computed no matter how firms make their investment decisions. If L and P are chosen optimally, one only needs to know the distribution of A and K to be able to compute TFP. One can show that TFP satisfies the following equation

$$\left(TFP^{\sigma-1}\right)^{\theta} K_s^{1-\theta} = \int \left(\tilde{A}_i^{\sigma-1}\right)^{\theta} K_i^{1-\theta} \propto \int A_i^{\theta} K_i^{1-\theta}$$

A special case worth mentioning is that if the conditional distribution of productivity/capital ratios, Z, is the same for every A, this reduces to

$$\left(TFP^{\sigma-1}\right) \propto \left(\int A_i\right) \left(\left[\frac{\mathbb{E}\left[\left(Z^{-1}\right)^{1-\theta}\right]}{\mathbb{E}\left[Z^{-1}\right]^{1-\theta}}\right]\right)^{\frac{1}{\theta}}$$

Here TFP can be written as two terms, one of which depends only on productivity parameters and one which depends on the distribution of Z's. One can use Jensen's inequality to show that this second term is less or equal to 1, with equality only if every firm has the same productivity/capital ratio. The actual level of Z's doesn't affect TFP, only the dispersion. A widening of the inaction region will leave the first term unchanged (up to entry and exit) but will decrease the second term as the firm's fan out across the inaction region.

2.3.3 Dynamics

Given a sequence of thresholds, the evolution of the distribution of firms over the state space can be described by the Kolmogorov forward equation and several boundary conditions. These can be derived by taking the limit of a discrete approximation to the state space. Firms enter at rate E_t, and all entering firms productivity A_0.

Let $m(a,k,t)$ be the density of firms at time t, where $a \equiv \ln A$ and $k \equiv \ln K$. In the interior of the inaction region, the distribution evolves according to the KFE:

$$m_t = -\chi m + \delta m_k - \mu m_a + \frac{\varepsilon^2}{2} m_{aa}$$

The fact that this is a partial differential equation in three variables makes it difficult to solve for the invariant distribution, let alone transition dynamics. However, it turns out that to characterize the objects of interest, it is not necessary to solve it directly.

Consider two economies with different parameterizations:

	1	2
Depreciation	δ	0
Drift of Productivity	μ	$\mu + \delta$
Arrival Rate of Death Shock	χ	$\chi + \delta$

and all other parameters the same. If both economies have the identical entry rates and realization of interest rate shocks, then the two economies will have identical total output, capital stock, investment and TFP. Furthermore firms in each economy will have the same investment and disinvestment thresholds[9].

This is useful because it is much easier to track the distribution of firms in an economy with no

[9] Consider two firms with different parameterizations but that experience the same shock processes (ω_t and r_t). It is easy to show that they will have the same investment thresholds and investment timing (conditional on not receiving the death shock). Let $A_j(t)$, $K_j(t)$, $Y_j(t)$, $dI_j^+(t)$, and $dI_j^-(t)$ be the productivity, capital stock, output, and net investment of firm with parameterization j. It is easy to show that if $\{I_1^{+*}, I_1^{-*}\}$ are the optimal investment and disinvestment policies for firms with parameterization 1, then $\{I_2^{+*}, I_2^{-*}\}$ defined so that $dI_2^{+*} = e^{\delta t} dI_1^{+*}$ and $dI_2^{-*} = e^{\delta t} dI_1^{-*}$ are optimal policies for firms with parameterization 2. Furthermore

$$\begin{aligned} A_2(t) &= e^{\delta t} A_1(t) \\ K_2(t) &= e^{\delta t} K_1(t) \\ Y_2(t) &= e^{\delta t} Y_1(t) \end{aligned}$$

Firm 2, is essentially a scaled version of firm 1.

Now consider a cohort of firms under parameterization 1 and a cohort under parameterization 2, both entering at time zero. Each cohort will have the same total output, capital stock, cumluative investment and disinvestment, and TFP.

depreciation: it reduces the number of variables with derivatives in the KFE

$$m_t = -(\chi + \delta) m - (\mu + \delta) m_a + \frac{\varepsilon^2}{2} m_{aa}$$

In particular, the KFE associated with the invariant distribution (with $m_t = 0$) is simply an ordinary differential equation.

If $\bar{z}(t)$ and $\underline{z}(t)$ are the logs of the thresholds at time t, we can write the additional boundary conditions as (note that $k = a - z$):

Upper Boundary:

$$\mu m(a, a - \bar{z}(t), t) = m_a \varepsilon^2 (a, a - \bar{z}(t), t) + m_k (a, a - \bar{z}(t), t) \frac{\varepsilon^2}{2}$$

Lower Boundary:

$$\mu m(a, a - \underline{z}(t), t) = m_a (a, a - \underline{z}(t), t) \varepsilon^2 + m_k (a, a - \underline{z}(t), t) \frac{\varepsilon^2}{2}$$

At the point of entry

$$\frac{\varepsilon^2}{2} \left(m(a_0, a_0 - \bar{z}(t), t) - \lim_{\xi \to 0} m(a_0 - \xi, a_0 - \bar{z}(t) - \xi, t) \right) = E_t$$

With these equations one can derive a closed form expression for the steady state distribution of firms and precisely describe the evolution of the distribution after a shock to the real interest rate.

If a set of thresholds with slopes $\bar{Z}$ and $\underline{Z}$ and entry at rate E have been in place for an arbitrarily long time, the distribution of firms will satisfy:

$$m^{ss}(a, k) = \begin{cases} \frac{E}{\varepsilon^2/2} \frac{1}{1 - e^{-\gamma x}} e^{\eta_N (a - a_0)} e^{-\gamma(\bar{z} - z)}, & K > K_0 \\ \frac{E}{\varepsilon^2/2} \frac{1}{e^{\gamma x} - 1} e^{\eta_P (a - a_0)} e^{\gamma(\bar{z} - z)}, & K < K_0 \end{cases}$$

where η_P and η_N are the positive and negative roots of $(\chi + \delta) = -(\mu + \delta)\eta + \frac{\varepsilon^2}{2}\eta^2$ respectively,

and $\gamma = \sqrt{\left(\frac{\mu+\delta}{\varepsilon^2/2}\right)^2 + 4\frac{\chi+\delta}{\varepsilon^2/2}}$. $m^{ss}(a,k)$ can be interpreted as the measure of firms at (a,k). For this to be a proper density, simply set $E = \chi$.

Among firms sufficiently above the entry threshold, the conditional distribution of productivity capital ratio, Z, is a truncated Pareto with exponent γ. The fact that this distribution is identical for each A is useful. In fact, as long as the the entry rate remains constant, the conditional distribution of Z will be identical for each productivity level A even when the thresholds vary over time (again, for A sufficiently above the entry point). This is helpful because it leads to a convenient approximation. Since the dynamics at the sector level are mostly dominated by the actions of large firms, the dynamics of firms at a given productivity level will be a very good approximation for the dynamics of the entire sector. I will analyze the dynamics of this conditional distribution.

2.3.4 Overshooting

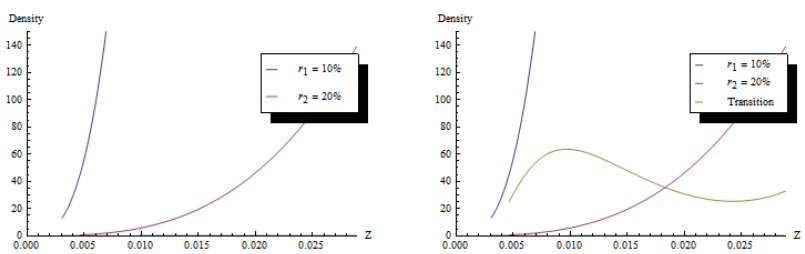
Depending on the parameters, the short run decline in aggregate TFP following a shock to the interest rate will be even larger than that implied by the widened (s,S) bands. If firms invest more frequently than they disinvest (i.e., depreciation and shocks to idiosyncratic productivity tend to push firms toward the investment threshold), the steady state distribution of firms will be clustered close to that threshold. Once the interest rate shock hits, the investment threshold shifts up. Productivity shocks and depreciation will push firms to fan out across the entire inaction. If the productivity shock lasted an arbitrarily long time, the distribution of firms would eventually settle into what would be a steady state distribution for the new thresholds, clustered close to the new investment threshold. TFP is decreasing in the cross-sectional dispersion of marginal products of capital. This dispersion will be higher when the distribution of firms is spread out across the state space than when it becomes clustered close to the new investment threshold.

As shown in section 2.3.3, one can compute the steady state distribution for any given set of thresholds (as long as the rate of entry is constant). Figure 8(a) shows the a cross section of invariant distribution (the distribution of Z conditioned on some productivity level A sufficiently above the entry point) for the thresholds associated with each regime. The distribution for the high interest rate regime can be interpreted as the outcome if the high interest regime, though expected

to be short, ended up lasting forever. The thresholds are farther apart for the high interest rate regime, but for both regimes the distribution is concentrated close to the upper threshold.

At the onset of an interest rate shock, the investment threshold will shift up. The firms that were closest to that threshold will spread out over the entire inaction region due to depreciation and positive productivity shocks. Figure 8(b) show the distribution when the interest rate shock has lasted a year. The distribution of Z's is clearly more spread out during this point in the transition than at either of the two endpoints. It follows that if the interest rate shock lasted an arbitrarily long time, TFP would be lowest when the distribution was most spread out, which would be at some point during the transition. After that point, TFP would rise.

Figure 8:



(a) (b)

2.3.5 Impulse Response

I now consider the experiment of an economy that has been in the low wage regime for a long time, but firms account for the possibility of a shift to a high wage regime. I compute the impulse response to a wage shock that is expected to be short-lived.

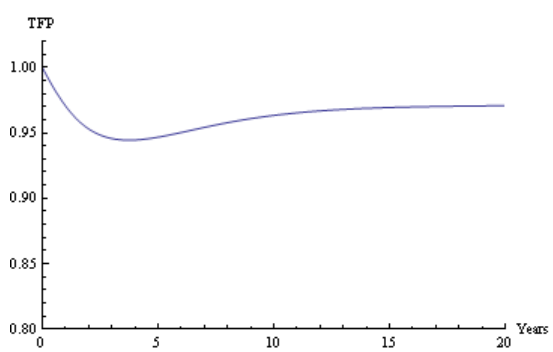
I calibrate the model with the following parameters

Variable	Description	Value
r_1	Low Interest Rate	10%
$\frac{\sigma}{\sigma-1}$	Markup	1.2
$1-\alpha$	Labor Share	.5
δ	Depreciation	0.1
χ	Exogenous Death	.05
p^+	Capital Price	1
p^-	Resale Price	0.6
μ	Drift	0.02
ε	Volatility	0.15

These give the following moments:

Moment	Value
Investment/Capital Stock	0.16
% of plants investing	57%
% disinvesting	1%

I compute the TFP relative to the initial steady state. Over two years, one can see a fairly large decline in TFP, and the impact of the overshoot is evident.



I conclude that an interest rate shock can have a significant impact on aggregate TFP.

2.4 Real Wages

The effect of a change in real wages depends on how flexibly firms can adjust employment. Consider first the model presented above, in which firms flexibly and fully adjust employment. An increased wage will reduce employment, and lower the effective productivity of capital. In this sense, an increased wage would have the same effect on the (s,S) bands as negative productivity shock. An short-lived wage increase would therefore compound the of a short-lived interest rate spike on the (s,S) bands.

To the extent that firms do not adjust employment (e.g., due to hiring frictions), there will be no additional effect on investment. However, aggregate TFP will still decline as the dispersion of marginal product of labor increases. The mechanism is almost identical to that described in the previous section. There are two differences from the shock to interest rates shock. First, while the interest rate helps determine the option value of waiting, a change in the wage has no such effect. Second, the wage may be a determinant of the cost of hiring or firing, whereas the interest rate played no role in the purchase price of capital. The cost of firing a worker can include severance pay, which is tied to the real wage. Similarly, the cost of recruiting and training a worker may be denominated in time, in which case it the monetary cost of hiring would be proportional to the wage. If the hiring and firing costs are proportional to the wage, during a short-lived wage spike a firm would be even more reluctant to adjust employment because it could simply avoid these elevated costs once the shock dissipated. This means that during a wage spike, the (s,S) bands would expand even more[10].

In summary, shocks to the real wage are likely to reduce aggregate TFP. However, this may occur through increased cross-sectional dispersion of either marginal products of capital or of marginal products of labor, depending on the flexibility of employment adjustments.

[10] Cutting in the other direction, particularly for Chile in the early 1980's, is the fact that changes in the real wage are more likely to be anticipated than changes in the interest rate, as they are were often part of legislation. Firms may make adjustments just before an anticipated wage increase to avoid the larger adjustment costs.

3 Reduced Form Empirics

3.1 Data Description

I use data from the annual census of the manufacturing sector in conducted by Chile's Instituto Nacional de Estadistica spanning the year 1979 to 1996. The survey covers all manufacturing plants with at least 10 employees and has been used by numerous other studies. The dataset is an unbalanced panel with nearly unique 5000 observations each year. The survey collects data on revenue, blue and white collar employment and wages, and intermediate inputs. Plants report investment, sales, and depreciation of several types of capital. Entry and exit from the dataset does not correspond to actual entry and exit; a firm that shrinks to nine employees would not show up in the data. Indeed, some firms disappear in one year only to reappear a year or two later. The version of the data that I have contains no information linking plants to firms, though the vast majority of plants belong to single plant firms[11].

3.2 Average Revenue Product of Capital and Labor

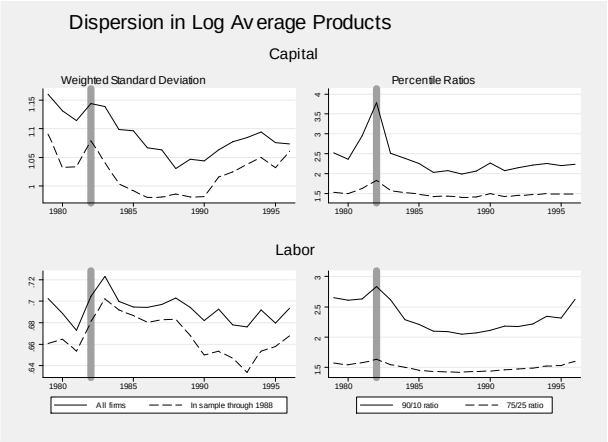
A key prediction is that the distribution of the productivity/capital, $Z = A/K$ ratio gradually becomes more dispersed while the interest rate is high. Again, the empirical analog of Z is the average revenue product of capital, $\log Z$ is a linear function of $\ln\left(\frac{PY}{K}\right)$. To get a sense of the evolution of the distribution of Z, we can look at the distribution of $\frac{PY}{K}$, the average revenue product of capital. If labor rigidities are important, we should also expect to see increased dispersion in the average revenue product of labor.

I report here the evolution of the cross sectional dispersion of $\ln\left(\frac{PY}{K}\right)$ and $\ln\left(\frac{PY}{WL}\right)$. The measure of capital is a weighted average of the different types of capital, weighted by their respective user costs. If workers are heterogeneous, there may be skill differences across plants. Wage paid is a proxy for the skill of a worker. Analyzing the logs of these objects is convenient because changes in aggregate price levels will not affect the measures of dispersion.

The first plot in figure 9 shows the standard deviation of log of the average product of capital. The solid line contains all firms in the sample, while the dashed line contains only the firms that

[11] Hsieh and Parker (2007), using information provided by the INE, report that most plants in the dataset are themselves firms. In 1984, approximately 350 plants belonged to multi-plant firms, under 10%.

Figure 9:



remained in the sample over the first ten years. The model speaks to changes in the dispersion keeping the composition of firms constant, so I view the dashed line as more empirically relevant. This plot clearly shows a big increase in the standard deviation of log average product of capital.

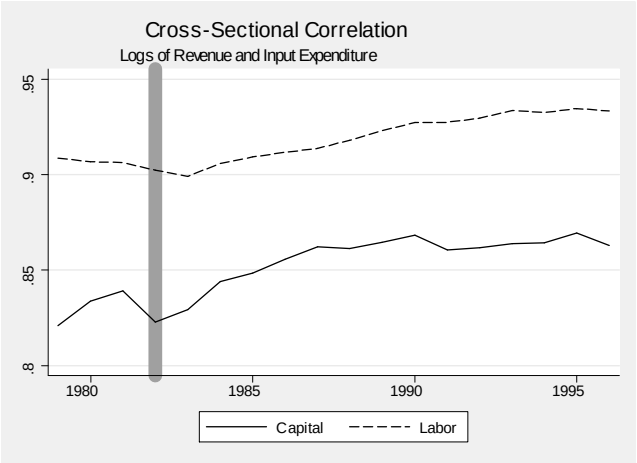
The standard deviation may be sensitive to outliers, so I plot two additional measures of dispersion, the 90-10 ratio and the 75-25 ratio. These measure also point to rising dispersion in 1981, peaking in 1982.

I plot similar figures for labor. The weighted standard deviation of the APL is rising in 1982, though it is not clear that it is especially high while the wage is high. In contrast, the 90-10 and 75-25 ratios show increased dispersion

Another Prediction of the model is that the allocation of resources becomes less efficient. One measure of the efficiency of the allocation of resources across firms is the cross-sectional correlation of capital and revenue. Bartelsman et al. (2009) use a similar measure, the correlation of labor and firm level solow residual. I prefer to use output to the solow residual because it relies less on modeling assumptions and because using the solow residual builds in artificial negative correlation if there is mismeasurement of capital. Petrin and Levinsohn (2008) rightly criticize the use of this type of correlation as a measure of the efficiency of the allocation of resources because it is not motivated by theory. The same critique applies here. The decomposition in section 4 takes a more structured approach to the efficiency of the allocation of resources.

Figure 10 shows the correlation of log revenue with log capital expenditures and of log revenue with log labor expenditures. By this measure, there was a worsening of the allocation of capital

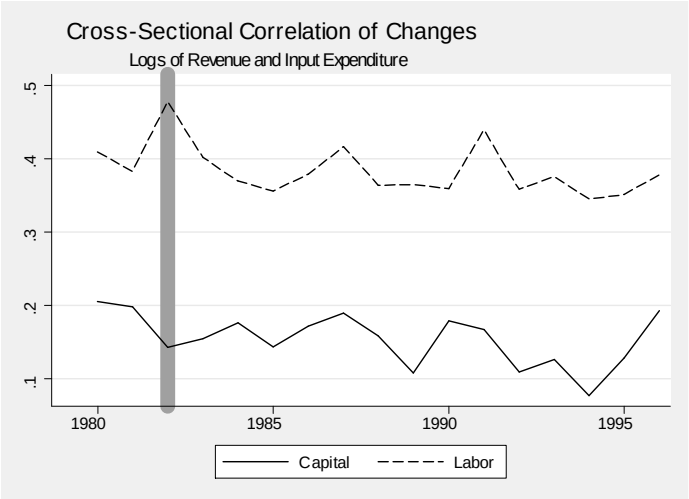
Figure 10:



across firms in 1982. The allocation of labor gets slightly worse, and continues to decline in 1983, though this change is relatively small.

We can also analyze the pace of reallocation by looking at the correlation of changes in revenue with changes in inputs. A dip in this correlation means that reallocation of resources slows down. There is certainly a decrease in the correlation between changes in revenue and changes in capital

Figure 11:



in 1982, but it is hard to argue that the correlation is especially low. The correlation between revenue and labor rises as many workers were laid off in 1982.

4 Business Cycle Accounting at the Micro Level

To quantify the impact of changes in the dispersion of average products on the decline in TFP, I develop an accounting procedure that allows me to separate changes in TFP into changes in firm level productivity and changes in the distribution of average products.

The decomposition is interpretable for economists doing business cycle accounting at the aggregate level using a representative firm. The Solow residual is decomposed into the various parts described above, but there is also a role for wedges in the first order conditions of the representative firm. If the wedge at each individual firm is equal to the wedge for the representative firm, the Solow residual will be equal to the aggregate of firm level productivities, with no role for plant level dispersion. In this sense, the decomposition bridges the gap between the approaches of Chari et al. (2007) and Hsieh and Klenow (2009).

The decomposition uses only nominal values of revenue and expenditures that are directly reported in the survey (except for expenditure on capital services). In an inflationary environment such as Chile in the early 80's, prices could be measured poorly, so using price deflators convert expenditures to quantities can introduce more noise into measurements. In addition, using price deflators that are common for an entire industry[12] can mask differences in output and input prices facing different firms. These differences are precisely what the decomposition is attempting to uncover. Using expenditures on inputs also helps to control for differences in input quality. For example, the wage bill can be used to control for the quality of the workforce, whereas deflating by a common wage (or simply using the number employees) could conflate quantity and price differences.

A notable feature of the Chilean plants is there are large, persistent differences across plants in capital intensity. For example, large plants are, on average, much more capital intensive than small plants. This could simply reflect the differences in technology across plants. Alternatively this may reflect persistent distortions faced by some plants. This matters because if the differing capital usage reflects differing technologies there is little to gain from reallocating resources across plants. I will consider two empirical specifications allowing for each of these extremes.

[12] Several other decompositions, including Petrin and Levinsohn (2008), use common price deflators for the industry.

4.1 The Decomposition

Consider an economy composed of many sectors. Each sector is composed of plants that produce differentiated products that can be combined into a sector aggregate[13]. These sector aggregates are then combined into a single aggregate good. Let Y_i be the output of firm i, $Y_s = \left(\sum_{i \in I_s} Y_i^{\frac{\sigma_s - 1}{\sigma_s}}\right)^{\frac{\sigma_s}{\sigma_s - 1}}$ be the quantity of sector aggregate for sector s, and $Y = \left(\sum_s Y_s^{\frac{\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\gamma-1}}$ be the quantity of the manufacturing aggregate. If P_i is the price of the good produced by firm i, then $P_s = \left(\sum_{i \in I_s} P_i^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ and $P = \left(\sum_s P_s^{1-\gamma}\right)^{\frac{1}{1-\gamma}}$ are ideal price indices for the sectoral good and aggregate good respectively[14]. Cost minimization implies that demand for good i is $Y_i = Y\left(\frac{P_s}{P}\right)^{-\gamma}\left(\frac{P_i}{P_s}\right)^{-\sigma}$.

Firm i has access to the production function $Y_i \leq A_i K_i^{\alpha_{Ki}} L_i^{\alpha_{Li}} M_i^{\alpha_{Mi}}$ and faces demand $Y_i \leq D_s P_i^{-\sigma_s}$. Define the wedges τ_{Ki}, τ_{Li},τ_{Mi} to satisfy the following equations[15]:

$$
\begin{aligned}
e^{\tau_{Ki}} &= \frac{\sigma - 1}{\sigma}\alpha_{Ki}\frac{P_i Y_i}{RK_i} \\
e^{\tau_{Li}} &= \frac{\sigma - 1}{\sigma}\alpha_{Li}\frac{P_i Y_i}{WL_i} \\
e^{\tau_{Mi}} &= \frac{\sigma - 1}{\sigma}\alpha_{Mi}\frac{P_i Y_i}{QM_i}
\end{aligned}
$$

These wedges measure deviations from optimization in a frictionless world. Whatever the reasons for the wedges, the firms actions will be observationally equivalent to a firm maximizing profits in a frictionless world subject to distortions:

$$
\pi_i = \max_{P_i, Y_i, K_i, L_i, M_i} P_i Y_i - e^{\tau_{Ki}} RK_i - e^{\tau_{Li}} WL_i - e^{\tau_{Mi}} QM_i
$$

[13] Some argue that declines in productivity are driven by reallocation across sectors. If this is the case, the decline should show up in increased dispersion of between sector wedges. It is useful to have an intermediate sector level so that dispersion in wedges between sectors is not conflated with dipersion in wedges within sectors.

[14] Throughout, variables with an i subscript are refer to firm i in sector s, variables with an s subscript refer to aggregates for sector s, and variables with no subscript refer to aggregates for all manufacturing plants. Time subscripts are omitted.

[15] These can be written a number of ways. There are four observable quantities ($P_{si}Y_{si}$, WL_{si}, RK_{si}, QM_{si}). The first order conditions are all linear combinations of the logs of these quantities, and there are only three linearly independent differences in these quantities. However, there is no unique way to assign meanings to the wedges - any linear combination will do just as well - so interpreting a particular wedge is perilous.

One can solve for the price set by the firm

$$P_i = \frac{1}{\lambda_i A_i} e^{\tau_i}$$

where $\tau_i \equiv \alpha_{Ki}\tau_{Ki} + \alpha_{Li}\tau_{Li} + \alpha_{Mi}\tau_{Mi}$ aggregates the effect of the three wedges and $\lambda_i \equiv \frac{\sigma_s - 1}{\sigma_s} \frac{\alpha_{Ki}^{\alpha_{Ki}} \alpha_{Li}^{\alpha_{Li}} \alpha_{Mi}^{\alpha_{Mi}}}{R^{\alpha_{Ki}} W^{\alpha_{Li}} Q^{\alpha_{Mi}}}$ reflects the impact of changes in input prices.

Sector

At the sector level, an economist may define a representative firm with technology $Y = TFP \times K^{\alpha_s} L^{\alpha_L} M^{\alpha_M}$. One can define wedges at the sector level, τ_{Ks}, τ_{Ls}, τ_{Ms} to satisfy the equations

$$\begin{aligned} e^{\tau_{Ks}} &= \frac{\sigma_s - 1}{\sigma_s} \alpha_{Ks} \frac{P_s Y_s}{RK_s} \\ e^{\tau_{Ls}} &= \frac{\sigma_s - 1}{\sigma_s} \alpha_{Ls} \frac{P_s Y_s}{WL_s} \\ e^{\tau_{Ms}} &= \frac{\sigma_s - 1}{\sigma_s} \alpha_{Ms} \frac{P_s Y_s}{QM_s} \end{aligned}$$

Again, this can be thought of as arising from the optimization problem of a firm facing demand $Y_s = DP_s^{-\sigma_s}$ and several distortions:

$$\pi_s = \max_{P_s, Y_s, K_s, L_s, M_s} P_s Y_s - e^{\tau_{Ks}} RK_s - e^{\tau_{Ls}} WL_s - e^{\tau_{Ms}} QM_s$$

The price set by the firm is

$$P_s = \frac{e^{-\tau_s}}{\lambda_s TFP_s}$$

where $\tau_s \equiv \alpha_{Ks}\tau_{Ks} + \alpha_{Ls}\tau_{Ls} + \alpha_{Ms}\tau_{Ms}$ and $\lambda_s = \frac{\sigma_s - 1}{\sigma_s} \frac{\alpha_{Ks}^{\alpha_{Ks}} \alpha_{Ls}^{\alpha_{Ls}} \alpha_{Ms}^{\alpha_{Ms}}}{R^{\alpha_{Ks}} W^{\alpha_{Ls}} Q^{\alpha_{Ms}}}$ are defined in the same way as τ_i and λ_i. We can equate the two expressions for the price level and $P_s^{1-\sigma_s} = \sum_i P_i^{1-\sigma_s}$ to get the following expression:

$$\lambda_s^{\sigma_s - 1} TFP_s^{\sigma_s - 1} = \sum_i \left(\frac{\lambda_i A_i}{\lambda_s} \right)^{\sigma_s - 1} e^{(\sigma_s - 1)(\tau_s - \tau_i)}$$

We now define the quantity A_s to satisfy $\lambda_s A_s \equiv \left(\sum_i \lambda_i A_i^{\sigma-1}\right)^{\frac{1}{\sigma-1}}$ to be what sector level TFP would be if all the wedges were zero. We can also define $B_i \equiv (\lambda_i A_i/\lambda_s A_s)^{\sigma_s-1}$ to be the revenue share of firm i if there were no wedges. With these definitions we can write a simple expression for a sector's TFP relative to a frictionless world,

$$TFP_s = A_s \left(\sum_i B_i e^{(\sigma_s-1)(\tau_s-\tau_i)}\right)^{\frac{1}{\sigma_{s-1}}}$$

There are two components. The first depends only on firm level technology parameters (and if factor intensities differ across firms, the common prices of the various inputs). The second term depends on the wedges across firms. If all wedges were equal to the industry wedge, this term would be zero. If there is a large amount of dispersion across firms in wedges, this term will be significantly below one and TFP in the industry will be smaller than it could be. In this sense, changes in the second term measure the contribution of dispersion to changes in sector level TFP.

To measure relative TFP, $\frac{TFP_s}{A_s}$, I need expressions for the wedges and for the frictionless shares, B_i. The wedges can be computed directly using the ratios of nominal revenue and payments to the factors of production. To compute the frictionless shares, note that we can use the demand system, $Y_i = D_s P_i^{-\sigma}$ to write output as $Y_i = D_s^{\frac{1}{1-\sigma}} (P_i Y_i)^{\frac{\sigma}{\sigma-1}}$. With this and the production function $A = Y/\left(K^{\alpha_{Ki}} L^{\alpha_{Li}} M^{\alpha_{Mi}}\right)$, we can get an expression for firm level productivity $\lambda_i A_i$:

$$\lambda_i A_i \propto \alpha_{Ki}^{\alpha_{Ki}} \alpha_{Li}^{\alpha_{Li}} \alpha_{Mi}^{\alpha_{Mi}} \frac{(P_i Y_i)^{\frac{\sigma_s}{\sigma_s-1}}}{(RK_i)^{\alpha_{Ki}} (WL_i)^{\alpha_{Li}} (QM_i)^{\alpha_{Mi}}}$$

which consequently gives an expression for $B_i = \frac{(\lambda_i A_i)^{\sigma_s-1}}{\sum_i (\lambda_i A_i)^{\sigma_s-1}}$. Conveniently this expression depends only on nominal revenue and payments to the factors of production.

Aggregate Level

One can extend this procedure to a representative firm at the aggregate level with the production function $Y \leq TFP \times K^{\alpha_K} L^{\alpha_L} M^{\alpha_M}$ and facing demand $Y \leq \bar{D} P^{-\sigma}$. Defining the wedges τ_K, τ_L, τ_M in the same manner, the optimal price will be $P = \frac{e^{\tau}}{\lambda TFP}$, with $\lambda = \frac{\sigma-1}{\sigma} \frac{\alpha_K^{\alpha_K} \alpha_L^{\alpha_L} \alpha_M^{\alpha_M}}{R^{\alpha_K} W^{\alpha_L} Q^{\alpha_M}}$ and $\tau \equiv \alpha_K \tau_K + \alpha_L \tau_L + \alpha_M \tau_M$. Again, using the relation $P^{1-\gamma} = \sum P_s^{1-\gamma}$ gives an expression for

TFP at the aggregate level in terms of the relative total factor productivity of each of the sectors:

$$\left(\frac{TFP}{A}\right)^{\gamma-1} = \sum B_s \left(\frac{TFP_s}{A_s}\right)^{\gamma-1} e^{(\gamma-1)(\tau-\tau_s)}$$

As above, we can define A (satisfying $(\lambda A)^{\gamma-1} \equiv \sum_s (\lambda_s A_s)^{\gamma-1}$) to be aggregate technology in a wedgeless world, and $B_s \equiv (\lambda_s A_s/\lambda A)^{\gamma-1}$ to be the revenue share of sector s. We can write TFP in the industry as

$$TFP = A \left\{ \sum_s B_s e^{(\gamma-1)(\tau-\tau_s)} \left(\sum_i B_i e^{(\sigma_s-1)(\tau_s-\tau_i)} \right)^{\frac{\gamma-1}{\sigma-1}} \right\}^{1/(\gamma-1)}$$

Here, we can see that TFP at the aggregate level depends on an aggregate of firm level technology parameters and a term that depends on wedges within and between sectors.

Interpreting the Representative Firm

If every firm's technology has the same factor intensities, the production function of the representative firm should have the same factor intensities. If individual firms differ in factor intensities, there is no obvious way to construct a representative firm that aggregates nicely. For example, consider a sector with two firms: firm 1 is capital intensive while firm 2 is labor intensive. If firm 1 has a large share of the market, the sector will look capital intensive (for the most efficient allocation) and vice versa. It is not clear what form the representative production function should take.

In principle, the decomposition can be done for any representative production function. However, a desirable feature is that if all firms in an industry have the same capital wedge τ_K^0, then the representative firm will have that same capital wedge τ_K^0. A consequence of this is that if every firm had wedges of 0, TFP_s would be equal to A_s. To guarantee these features, one can set choose factor intensities for the representative production function as the average factor intensities of each of the firms, weighted by revenue, for example $\alpha_{Ks} = \sum_i \alpha_K \frac{P_i Y_i}{P_s Y_s}$ and $\alpha_K = \sum_s \alpha_{Ks} \frac{P_s Y_s}{PY}$. Similarly, one must assume that the markup at the aggregate level is a weighted average of the markup of each sector, $\frac{\sigma-1}{\sigma} = \sum_s \frac{\sigma_s-1}{\sigma_s} \frac{P_s Y_s}{PY}$. See appendix C for details.

One unfortunate consequence is that the representative production function will change as the

sector's composition shifts. Again, this is not an issue when every firm has the same α_{Ki}.

4.2 Application to the Chilean Data

Firm Specific Parameters

In this section I adhere to the view that all long run differences in factor intensities reflect differences in underlying technology. To this end, I compute the parameters of each firm's production function by using the average cost share of each input over the years that the plant is in the sample. Similarly I compute the elasticity of substitution for each sector averaging the markup implied revenue/cost margin in each year.

While the plants provide data on revenue, the wage bill, and expenditures on intermediate materials, the expenditure on capital services must be imputed. Even after constructing a series for capital, one must still infer the user cost of capital. This depends on the depreciation rate, the interest rate, and the expected change in price of capital. I have no way of directly measuring the expected change in the price of capital, though capital price deflators show large swings in the prices of buildings, machinery, and vehicles during the contraction. If these changes are expected, there would be massive and unrealistic swings in capital's share of revenue (for some years, the user cost of certain types of capital would be negative). While this is not a completely innocuous simplification, I assume that all changes in real capital prices are unexpected. Thus variation in the user cost of capital will mostly reflect changes in the real interest rate.

Figure 12:

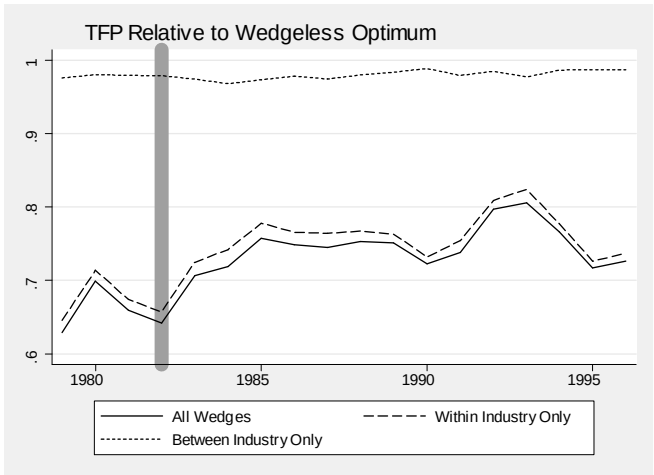
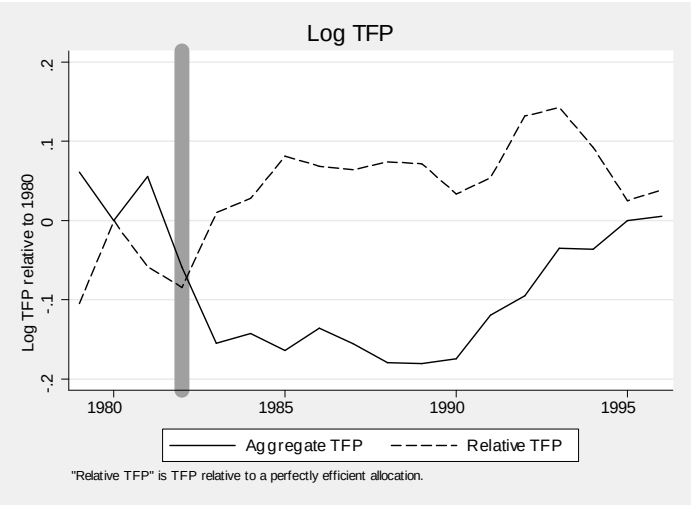


Figure 12 shows relative TFP, or $\frac{TFP}{A}$ for the economy. If changes in the wedges were the only

determinants of changes in TFP, than this line would perfectly track TFP (because A would be constant). The second line is what relative TFP if there were only dispersion in wedges within sectors. Lastly, the third line shows what relative TFP would be if there were only dispersion in wedges across sectors.

One can see that in 1981 and 1982, relative TFP falls significantly due to increased dispersion in wedges within sectors. It then rises abruptly in 1983. This is consistent with the prediction of the model that TFP slowly declines while the interest rate is high, and the rises once the interest rate shock dissipates.

Figure 13:

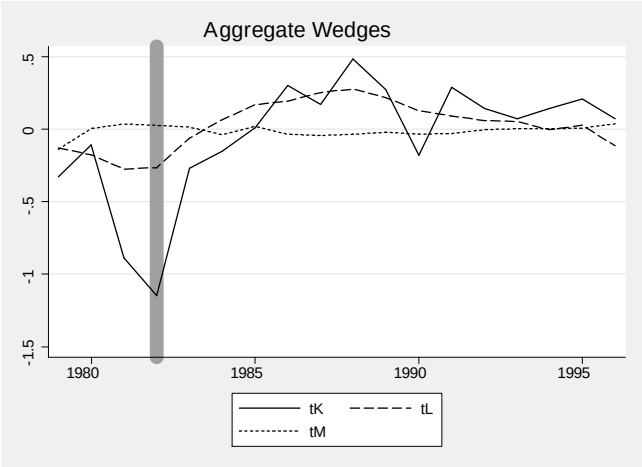


The next figure compares relative TFP to actual TFP that is computed using the representative production function for the manufacturing sector. Again, if all changes were due to changes in dispersion, these curves would line up perfectly. Changes in the gap between the two curves are the portions of changes in TFP that remains to be explained. It appears that this gap is substantial in 1981 and 1983, the years surrounding the recession. This indicates that factors other than dispersion of average products drive much of the changes in TFP.

It is also interesting to look at the wedges for the aggregate production function. It is clear that there is a substantial decrease in the capital wedge. From the perspective of distortive taxes, firms are holding so much capital (given the price) that the wedge is picking up what must be a large subsidy to capital.

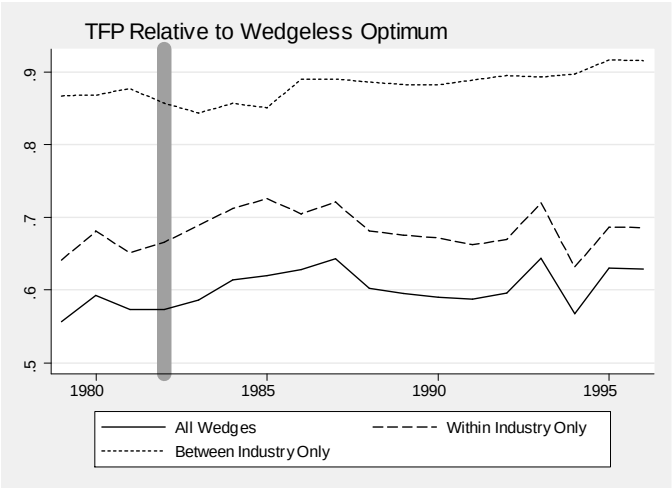
Constant Parameters

Figure 14:



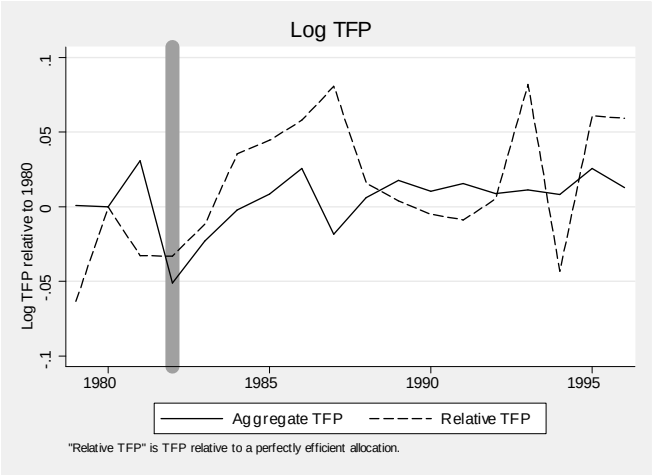
We can do the same decomposition, keeping the parameters the same for every firm. This conforms with the view that all long term differences in factor usage across plants stem from persistent distortions. The parameters of the production function are $\alpha_K = .15$, $\alpha_L = .35$, $\alpha_M = .5$, and $\sigma = 4$

Figure 15:



Again relative TFP drops from 1980 to 1982, though the contribution of within industry wedges actually increases TFP in 1982. It appears that dispersion in wedges across industries is decreases TFP in this period. Comparing the relative TFP to actual TFP, we can see that relative TFP matches a most of the decline in TFP from 1980 to 1982, but the timing of the decline does not match. In particular, the rise in actual TFP in 1981 does not match the decline in relative TFP.

Figure 16:



5 Conclusion

This paper makes several contributions. First I demonstrated that short-lived shocks to real interest rates can cause significant declines in TFP. In the presence of investment frictions, a spike in the real interest rate can exacerbate irreversibilities, widening the inaction region, leading to a less efficient allocation of capital. A key prediction is that while the interest rate is high, dispersion of the marginal product of capital gradually increases.

To test the model, I use data on Chilean Manufacturing firms and find that along some dimensions, the model fares fairly well. Dispersion in average products rises while the real interest rate is high, and this increased dispersion accounts for a substantial drop in aggregate TFP. Once the interest rate falls, dispersion falls as well.

While the increased dispersion causes a drop in TFP in 1981 and 1982, the exact timing of the changes in dispersion does not appear to be the main factor behind fluctuations in TFP. This result is somewhat sensitive to the exact empirical specification, as the decomposition relies heavily on the underlying model.

After this analysis, the time series in total factor productivity is still largely unexplained. It is possible that capacity utilization plays a role, and I would like to explore that more explicitly.

Another avenue for further research is to formally model a small open economy experiencing a spike to interest rates. While solving the model in general equilibrium is a nontrivial task, it is likely that it can be done using methods developed recently.

References

Abel, A. B. and J. C. Eberly, "Optimal Investment with Costly Reversibility," *Review of Economic Studies* 63 (1996), 581–93.

Banerjee, A. V. and E. Duflo, "Growth Theory through the Lens of Development Economics," in P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*volume 1 of *Handbook of Economic Growth*, chapter 7 (Elsevier, 2005), 473–552.

Bartelsman, E. J., J. C. Haltiwanger and S. Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocation and Selection," NBER Working Papers 15490, National Bureau of Economic Research, Inc, November 2009.

Bloom, N., "The Impact of Uncertainty Shocks," National Bureau of Economic Research, Inc, September 2007.

Bloom, N., S. Bond and J. V. Reenen, "Uncertainty and Investment Dynamics," *Review of Economic Studies* 74 (04 2007), 391–415.

Bloom, N., M. Floetotto and N. Jaimovich, "Really Uncertain Business Cycles," , 2009.

Campbell, J. R. and J. D. M. Fisher, "Aggregate Employment Fluctuations with Microeconomic Asymmetries," *American Economic Review* 90 (December 2000), 1323–1345.

Chari, V. V., P. J. Kehoe and E. R. McGrattan, "Sudden Stops and Output Drops," *American Economic Review* 95 (May 2005), 381–387.

———, "Business Cycle Accounting," *Econometrica* 75 (05 2007), 781–836.

Eisfeldt, A. L. and A. A. Rampini, "Capital reallocation and liquidity," *Journal of Monetary Economics* 53 (April 2006), 369–399.

Guo, X., J. Miao and E. Morellec, "Irreversible investment with regime shifts," *Journal of Economic Theory* 122 (May 2005), 37–59.

House, C. L., "Fixed Costs and Long-Lived Investments," NBER Working Papers 14402, National Bureau of Economic Research, Inc, October 2008.

Hsieh, C.-T. and P. Klenow, "Misallocation and Manufacturing TFP in China and India," Working Papers 09-04, Center for Economic Studies, U.S. Census Bureau, February 2009.

Hsieh, C.-T. and J. A. Parker, "Taxes and Growth in a Financially Underdeveloped Country: Evidence from the Chilean Investment Boom," *Economia* (2007).

Khan, A. and J. K. Thomas, "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?," *Journal of Monetary Economics* 50 (March 2003), 331–360.

Meza, F. and E. Quintin, "Factor Utilization and the Real Impact of Financial Crises," *Advances in Macroeconomics* 7 (2007), 1593–1593.

Neumeyer, P. A. and F. Perri, "Business cycles in emerging economies: the role of interest rates," *Journal of Monetary Economics* 52 (March 2005), 345–380.

Oberfield, E., "Irreversible Investment with Endogenous Entry and Exit," Mimeo, 2008.

Pavcnik, N., "Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants," *Review of Economic Studies* 69 (January 2002), 245–76.

Petrin, A. and J. Levinsohn, "Measuring Aggregate Productivity Growth Using Plant-Level Data," NBER Working Papers 11887, National Bureau of Economic Research, Inc, 2008.

Restuccia, D. and R. Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics* 11 (October 2008), 707–720.

Thomas, J. K., "Is Lumpy Investment Relevant for the Business Cycle?," *Journal of Political Economy* 110 (June 2002), 508–534.

Veracierto, M., "Firing Costs And Business Cycle Fluctuations," *International Economic Review* 49 (02 2008), 1–39.

Veracierto, M. L., "Plant-Level Irreversible Investment and Equilibrium Business Cycles," *American Economic Review* 92 (March 2002), 181–197.

A Constant Interest Rate: Optimality Conditions and Solution

A.1 Supercontact Conditions at Investment Thresholds

Let for any policy $\mathcal{P}$, we can write the solution to the HJB equations as

$$W(A,K|\mathcal{P}) = \frac{1}{1-\theta}\frac{1}{H(\theta)}A^{\theta}K^{1-\theta} - \frac{f}{r} + \sum_{i=1,2} c_i(K|\mathcal{P})A^{s_i}$$

where the constants of integration depend on the specific policy chosen. We first show that the thresholds are chosen so that the marginal effects on the constants of integration are zero.

Claim 1 *The optimal policy requires* $\frac{\partial c_i(K|\mathcal{P})}{\partial \bar{A}(K)} = \frac{\partial c_i(K|\mathcal{P})}{\partial \underline{A}(K)} = 0$ *for* $i = 1,2$ *and* $K > K_X$

Proof.

$$\begin{aligned} V(A,K) &= \max_{\mathcal{P}} W(A,K|\mathcal{P}) \\ &= \max_{\mathcal{P}} \frac{1}{1-\theta}\frac{1}{H(\theta)}A^{\theta}K^{1-\theta} - \frac{w}{r}K - \frac{f}{r} + \sum_{i=1,2} c_i(K|\mathcal{P})A^{s_i} \end{aligned}$$

For any K, the first order condition with respect to $\underline{A}(K)$ is $\sum_{i=1,2}\frac{\partial c_i(K|\mathcal{P})}{\partial \underline{A}(K)}A^{s_i} = 0$. By the principle of optimality, this holds $\forall A \in \left(\underline{A}(K), \bar{A}(K)\right)$. The only values of $\frac{\partial c_i(K|\mathcal{P})}{\partial \underline{A}(K)}$ that would satisfy these conditions are $\frac{\partial c_i(K|\mathcal{P})}{\partial \underline{A}(K)} = 0$. By an identical argument, $\frac{\partial c_i(K|\mathcal{P})}{\partial \bar{A}(K)} = 0$. ■

Claim 2 *The optimal thresholds satisfy* $W_{AK}\left(\bar{A}(K), K\right) = W_{AK}\left(\underline{A}(K), K\right) = 0$ *for* $K > K_X$

Proof. For each K we know that for any thresholds $\left(\bar{A}(K), \underline{A}(K)\right)$, the smooth pasting condition is holds

$$\begin{aligned} p^{+} &= W_K\left(\bar{A}(K), K|\mathcal{P}\right) \\ &= \frac{1}{H(\theta)}\bar{A}(K)^{\theta}K^{-\theta} + \sum_{i=1,2}\frac{\partial c_i(K|\mathcal{P})}{\partial K}\bar{A}(K)^{s_i} \end{aligned}$$

Differentiating implicitly with respect to $\bar{A}(K)$ gives (noting that the constants of integration depend on $\bar{A}$)

$$0 = \theta\frac{1}{H(\theta)}\bar{A}(K)^{\theta-1}K^{-\theta} + \sum_{i=1,2} s_i\frac{\partial c_i(K|\mathcal{P})}{\partial K}\bar{A}(K)^{s_i-1} + \sum_{i=1,2} s_i\frac{\partial^2 c_i(K|\mathcal{P})}{\partial \bar{A}(K)\,\partial K}\bar{A}(K)^{s_i}$$

Using Claim 1, we have

$$\frac{\partial^2 c_i\left(K|\mathcal{P}\right)}{\partial \bar{A}\left(K\right)\partial K}=\frac{\partial}{\partial K}\left(\frac{\partial c_i\left(K|\mathcal{P}\right)}{\partial \bar{A}\left(K\right)}\right)=\frac{\partial 0}{\partial K}=0$$

This delivers the supercontact conditions (the derivation for the lower threshold $\underline{A}\left(K\right)$ is identical)

$$W_{AK}\left(\bar{A}\left(K\right),K\right)=\theta\frac{1}{H\left(\theta\right)}\bar{A}\left(K\right)^{\theta-1}K^{-\theta}+\sum_{i=1,2}s_i\frac{\partial c_i\left(K|\mathcal{P}\right)}{\partial K}\bar{A}\left(K\right)^{s_i-1}=0$$

■

A.2 Solving for the thresholds

We first prove proposition 1.

Proposition 2 *For $K>K_X$, the upper and lower thresholds take the form $\bar{A}\left(K\right)=\bar{Z}K$ and $\underline{A}\left(K\right)=\underline{Z}K$*

Proof. Let $c_i\left(K\right)=\tilde{c}_i'\left(K\right)K^{s_i}$, $\bar{A}\left(K\right)=\bar{Z}\left(K\right)/K$, and $\underline{A}\left(K\right)=\underline{Z}\left(K\right)/K$. Substitute these into the smooth pasting and supercontact conditions to give the four equations

$$\begin{aligned}
p^+ &= \frac{1}{H\left(\theta\right)}\bar{Z}\left(K\right)^{\theta}+\sum_{i=1,2}c_i\left(K\right)\bar{Z}\left(K\right)^{s_i}\\
p^- &= \frac{1}{H\left(\theta\right)}\underline{Z}\left(K\right)^{\theta}+\sum_{i=1,2}c_i\left(K\right)\underline{Z}\left(K\right)^{s_i}
\end{aligned}$$

$$\begin{aligned}
0 &= \theta\frac{1}{H\left(\theta\right)}\bar{Z}\left(K\right)^{\theta}+\sum_{i=1,2}s_ic_i\left(K\right)\bar{Z}\left(K\right)^{s_i}\\
0 &= \theta\frac{1}{H\left(\theta\right)}\underline{Z}\left(K\right)^{\theta}+\sum_{i=1,2}s_ic_i\left(K\right)\underline{Z}\left(K\right)^{s_i}
\end{aligned}$$

For each K there are four equations in four unknowns, but note that the solution does not depend on K. We can therefore write $\bar{Z},\underline{Z}$, and c_i as constants. ■

The four equations can now be written as:

$$\begin{aligned} p^{+} &= \frac{1}{H(\theta)}\bar{Z}^{\theta} + \sum_{i=1,2} c_i \bar{Z}^{s_i} \\ 0 &= \theta \frac{1}{H(\theta)}\bar{Z}^{\theta} + \sum_{i=1,2} s_i c_i \bar{Z}^{s_i} \end{aligned}$$

$$\begin{aligned} p^{-} &= \frac{1}{H(\theta)}\underline{Z}^{\theta} + \sum_{i=1,2} c_i \underline{Z}^{s_i} \\ 0 &= \theta \frac{1}{H(\theta)}\underline{Z}^{\theta} + \sum_{i=1,2} s_i c_i \underline{Z}^{s_i} \end{aligned}$$

The smooth pasting and supercontact conditions for the upper threshold $\bar{Z}$ give

$$\begin{aligned} c_1 \bar{Z}^{s_1} &= \frac{s_2}{s_2 - s_1}\left[p^{+} - \left(\frac{s_2 - \theta}{s_2}\right)\left(\frac{1}{H(\theta)}\bar{Z}^{\theta}\right)\right] \\ c_2 \bar{Z}^{s_2} &= \frac{s_1}{s_1 - s_2}\left[p^{+} - \left(\frac{s_1 - \theta}{s_1}\right)\left(\frac{1}{H(\theta)}\bar{Z}^{\theta}\right)\right] \end{aligned}$$

Similarly the smooth pasting and supercontact conditions for the lower threshold give

$$\begin{aligned} c_1 \underline{Z}^{s_1} &= \frac{s_2}{s_2 - s_1}\left[p^{-} - \left(\frac{s_2 - \theta}{s_2}\right)\left(\frac{1}{H(\theta)}\underline{Z}^{\theta}\right)\right] \\ c_2 \underline{Z}^{s_2} &= \frac{s_1}{s_1 - s_2}\left[p^{-} - \left(\frac{s_1 - \theta}{s_1}\right)\left(\frac{1}{H(\theta)}\underline{Z}^{\theta}\right)\right] \end{aligned}$$

Eliminating the constants of integration gives

$$\begin{aligned} \left[p^{+} - \left(\frac{s_2 - \theta}{s_2}\right)\left(\frac{1}{H(\theta)}\bar{Z}^{\theta}\right)\right]\bar{Z}^{-s_1} &= \left[p^{-} - \left(\frac{s_2 - \theta}{s_2}\right)\left(\frac{1}{H(\theta)}\underline{Z}^{\theta}\right)\right]\underline{Z}^{-s_1} \\ \left[p^{+} - \left(\frac{s_1 - \theta}{s_1}\right)\left(\frac{1}{H(\theta)}\bar{Z}^{\theta}\right)\right]\bar{Z}^{-s_2} &= \left[p^{-} - \left(\frac{s_1 - \theta}{s_1}\right)\left(\frac{1}{H(\theta)}\underline{Z}^{\theta}\right)\right]\underline{Z}^{-s_2} \end{aligned}$$

and making the substitutions $X = \bar{Z}/\underline{Z}$ and $B = \frac{p^{+}}{p^{-}}$ and $\varphi = \frac{1}{rp^{+}}\bar{Z}^{\theta}$ and noting that $\frac{H(\theta)}{r} =$

$\left(\frac{s_1-\theta}{s_1}\right)\left(\frac{s_2-\theta}{s_2}\right)$ gives

$$\begin{aligned}\left[1-\frac{s_1}{s_1-\theta}\varphi\right] &= \left[B^{-1}-\frac{s_1}{s_1-\theta}\varphi X^{-\alpha}\right]X^{s_1}\\ \left[1-\frac{s_2}{s_2-\theta}\varphi\right] &= \left[B^{-1}-\frac{s_2}{s_2-\theta}\varphi X^{-\alpha}\right]X^{s_2}\end{aligned}$$

We can solve for each equation for φ to yield a single equation for

$$\frac{s_1-\theta}{s_1}\frac{X^{-s_1}-B^{-1}}{X^{-s_1}-X^{-\theta}}=\varphi=\frac{s_2-\theta}{s_2}\frac{X^{-s_2}-B^{-1}}{X^{-s_2}-X^{-\theta}} \tag{5}$$

Let $G(X,B,s)=\frac{s-\theta}{s}\frac{X^s-B}{X^s-X^\theta}$. we therefore have a single equation in one unknown which can be written in either of two ways

$$\begin{aligned}G\left(X^{-1},B^{-1},s_1\right) &= G\left(X^{-1},B^{-1},s_2\right)\\ G\left(X,B,s_1\right) &= G\left(X,B,s_2\right)\end{aligned}$$

From (5) we can also derive

$$\begin{aligned}\bar{Z}^\theta &= rp^+G\left(X^{-1},B^{-1},s_1\right)\\ \underline{Z}^\theta &= rp^-G\left(X,B,s_1\right)\end{aligned}$$

Lastly we have

$$\begin{aligned}c_1\underline{Z}^{s_1} &= \frac{s_2}{s_2-s_1}\left[p^- - \left(\frac{s_2-\theta}{s_2}\right)\frac{1}{H(\theta)}\underline{Z}^\theta\right]\\ c_2\underline{Z}^{s_2} &= \frac{s_1}{s_1-s_2}\left[p^- - \left(\frac{s_1-\theta}{s_1}\right)\frac{1}{H(\theta)}\underline{Z}^\theta\right]\end{aligned}$$

A.3 Comparative Statics

First I show that an increase in r raises s_1 and lowers s_2. In other words the spread between s_1 and s_2 increases. This seems intuitive because the product of the two roots is $-\frac{r}{\varepsilon^2/2}$.

Claim 3 $\frac{\partial s_1}{\partial r} = -\frac{\partial s_2}{\partial r} > 0$

Proof. $s_1 > 0 > s_2$ are the roots of $H(s) = r - \mu s - \frac{\varepsilon^2}{2} s(s-1) = 0$ with the formula $s_1, s_2 = \left(\frac{1-\frac{\mu}{\varepsilon^2/2}}{2}\right) \pm \sqrt{\left(\frac{1-\frac{\mu}{\varepsilon^2/2}}{2}\right)^2 + \frac{r}{\varepsilon^2/2}}$ ■

We now turn to the determinants of X. We first define the function $F(X, B, s_1, s_2) \equiv G(X, B, s_1) - G(X, B, s_2)$. X satisfies $F(X, B, s_1, s_2) = 0$.

Claim 4 $F_X(X, B, s_1, s_2) > 0$

Proof. Let $\phi(s) \equiv \frac{s}{X^s - 1}$. $\phi(s) > 0$ and $\phi'(s) < 0$. I will use the fact that

$$\begin{aligned} B - X^\theta G(X, B, s_2) &= B - \frac{s_2 - \theta}{s_2} \frac{X^{s_2} - B}{X^{s_2 - \theta} - 1} = B - \frac{X^s - B}{X^s - 1} \frac{\phi(s_2 - \theta)}{\phi(s_2)} \\ &< B - \frac{X^{s_2} - B}{X^{s_2} - 1} = \frac{B - 1}{1 - X^{-s_2}} < 0 \end{aligned}$$

and that $BX^{-\theta} - G(X, B, s_1) = BX^{-\theta} - G(X, B, s_2)$

$$\begin{aligned} G &= \left(1 - \frac{\theta}{s}\right) \frac{X^s - B}{X^s - X^\theta} \\ XG_X &= \left(1 - \frac{\theta}{s}\right) \frac{sX^s}{X^s - X^\theta} - \left(\frac{sX^s - \theta X^\theta}{X^s - X^\theta}\right) G \\ &= s\left(1 - \frac{\theta}{s}\right) \frac{X^s - B}{X^s - X^\theta} + \left(1 - \frac{\theta}{s}\right) s \frac{B}{X^s - X^\theta} - \left(s + \frac{(s-\theta) X^\theta}{X^s - X^\theta}\right) G \\ &= \frac{(s-\theta) B}{X^s - X^\theta} - \left(\frac{(s-\theta) X^\theta}{X^s - X^\theta}\right) G \\ &= \phi(s - \theta) \left[BX^{-\theta} - G\right] \end{aligned}$$

Putting these together we have

$$\begin{aligned} XF_X &= XG_X(X,B,s_1) - XG_X(X,B,s_2) \\ &= [\phi(s_1-\theta) - \phi(s_2-\theta)]\left[BX^{-\theta} - G\right] \\ &> 0 \end{aligned}$$

■

Claim 5 *$F_{s_1}(X,B,s_1,s_2) > 0$ and $F_{s_2}(X,B,s_1,s_2) < 0$*

Proof. I will proceed to show that $G_s(X,B,s) > 0$ when evaluated at $s \in (s_1, s_2)$. First note that $\eta\left(\frac{\log\eta}{\eta-1}\right)^2 < 1$ that implies that $\frac{\phi(s-\alpha)}{\phi(s)}$ is increasing in s as follows

$$\begin{aligned} \frac{\partial}{\partial s}\log\left(\frac{\phi(s-\alpha)}{\phi(s)}\right) &= \int_s^{s-\alpha} \frac{\partial^2}{\partial y^2}\log\phi(y)\,dy \\ &= \int_s^{s-\alpha} \frac{1}{y^2}\left(\frac{(X^{-y})(\log X^{-y})^2}{(1-X^{-y})^2} - 1\right)dy \\ &> 0 \end{aligned}$$

Second, for $B > 1$ note that $X^{s_1} > B > 1 > X^{s_2}$. We can see that the quantity $\frac{X^s-B}{X^s-1}$ is increasing in s as follows

$$\frac{\partial}{\partial s}\log\left(\frac{X^s-B}{X^s-1}\right) = \frac{B-1}{(X^s-B)(X^s-1)}X^s\log X > 0$$

Lastly, we can rewrite G as $G(X,B,s) = \frac{s-\theta}{s}\left(\frac{X^s-B}{X^s-X^\theta}\right) = \frac{1}{X^\theta}\frac{\phi(s-\theta)}{\phi(s)}\frac{X^s-B}{X^s-1}$. Since both $\frac{\phi(s-\theta)}{\phi(s)}$ and $\frac{X^s-B}{X^s-1}$ are increasing in s, G must also be increasing in s. ■

Putting claims together we can evaluate effect of r on X. Using the above claims, we have

$$\frac{\partial X}{\partial r} = -\frac{F_{s_1}\frac{\partial s_1}{\partial r} - F_{s_2}\frac{\partial s_2}{\partial r}}{F_X} < 0$$

A.4 Deriving Optimality Condition for the Exit Point

In proposition 1 we showed that the functions $\tilde{c}_i$ satisfy the differential equations $\tilde{c}_i'(K)K^{s_i} = c_i$. The numbers c_i are known with formulas given in Appendix A.2. We can solve each differential

equation to give

$$\tilde{c}_i(K) = \frac{c_i}{1-s_i} K^{1-s_i} + C_i$$

where C_i are new constants of integration. Since $s_1 > 0$, $\lim_{A\to\infty} C_1 A^{s_1}$ will diverge unless $C_1 = 0$. We can therefore write the value function as

$$W(A,K) = \frac{1}{1-\theta}\frac{1}{H(\theta)} A^{\theta} K^{1-\theta} - \frac{f}{r} + \sum_{i=1,2} \frac{c_i}{1-s_i} A^{s_i} K^{1-s_i} + C_2 A^{s_2} \tag{6}$$

where the only unknown is C_2. This will depend on the level of productivity at which the option to shut down is exercised.

Given optimal disinvestment thresholds, we can write the value of function as in (6)

$$W(A,K|\mathcal{P}) = \frac{1}{1-\theta}\frac{1}{H(\theta)} A^{\theta} K^{1-\theta} - \frac{f}{r} + \sum_{i=1,2} \frac{c_i}{1-s_i} A^{s_i} K^{1-s_i} + C_2(\mathcal{P}) A^{s_2} \tag{7}$$

The firm will choose an exit policy that maximizes value, giving the condition $\frac{\partial C_2(\mathcal{P})}{\partial A_X} = 0$. For any exit policy in which the firm shuts down an sells its remaining capital we have $W(A_X, K|\mathcal{P}) - p^- K = 0$ Differentiating this completely with respect to K gives

$$W_A(A_X,K) = \frac{1}{1-\theta}\frac{1}{H(\theta)} \theta A_X^{\theta-1} K^{1-\theta} + \sum_{i=1,2} \frac{c_i}{1-s_i} s_i A_X^{s_i-1} K^{1-s_i} + s_2 C_2(\mathcal{P}) A_X^{s_2-1} = 0$$

If we write $A_X = \underline{A}(K_X)$, we differentiate $W(\underline{A}(K_X), K_X) = p^- K_X$ implicitly with respect to K_X to get

$$W_K(A_X,K_X) + W_A(A_X,K_X)\,\underline{A}'(K_X) = p^-$$

Using the fact that $W_A(A_X,K_X) = 0$, we get $W_K(A_X,K_X) = p^-$

These equations give the following

$$p^- = \frac{W(A_X,K_X)}{K_X} = \frac{1}{1-\theta}\frac{1}{H(\theta)} \underline{Z}^{\theta} - \frac{f}{rK_X} + \sum_{i=1,2} \frac{c_i}{1-s_i} \underline{Z}^{s_i} + \frac{C_2(\mathcal{P})}{K_X} A_X^{s_2}$$

$$0 = \frac{A_X}{K_X} W_A(A_X,K) = \frac{1}{1-\theta}\frac{1}{H(\theta)} \theta \underline{Z}^{\theta} + \sum_{i=1,2} \frac{c_i}{1-s_i} s_i \underline{Z}^{s_i} + s_2 \frac{C_2(\mathcal{P})}{K_X} A_X^{s_2} = 0$$

We can solve the second equation for $\frac{c_2}{1-s_2}\underline{Z}^{s_2} + \frac{C_2(\mathcal{P})}{K_X}A_X^{s_2}$ and plug it into the first equation to get

$$\begin{aligned}\frac{f}{rK_X} &= -p^- + \frac{1}{1-\theta}\frac{1}{H(\theta)}\underline{Z}^\theta + \frac{c_1}{1-s_1}\underline{Z}^{s_1} - \frac{1}{s_2}\left(\frac{1}{1-\theta}\frac{1}{H(\theta)}\theta\underline{Z}^\theta + \frac{c_1}{1-s_1}s_1\underline{Z}^{s_1}\right) \\ &= -p^- + \left(1-\frac{\theta}{s_2}\right)\frac{1}{1-\theta}\frac{1}{H(\theta)}\underline{Z}^\theta + \left(1-\frac{s_1}{s_2}\right)\frac{c_1}{1-s_1}\underline{Z}^{s_1}\end{aligned}$$

We can use the relation $c_1\underline{Z}^{s_1} = \frac{s_2}{s_2-s_1}\left[p^- - \left(\frac{s_2-\theta}{s_2}\right)\frac{1}{H(\theta)}\underline{Z}^\theta\right]$ and $\underline{Z}^\theta = rp^-G(X,B,s_1)$, we can solve for the expressions given in the text.

B Solving for Optimal Policy with Two Regimes

Abstracting away from exit, let $V^i(Z,K)$ be the value of of being in state i. When in the inaction region of region i, the value function satisfies the HJB equation

$$r_iV^i(A,K) = \frac{1}{1-\theta}A^\theta K^{1-\theta} - \delta KV_K^i + \mu AV_A^i + \frac{\varepsilon^2}{2}A^2V_{AA} + \lambda_i\left(V^{-i}(A,K) - V^i(A,K)\right) - \chi V^i(A,K)$$

We take advantage of the homogeneity of the value function to define the function $v_i(Z) = V_K^i(ZK,K)$ as the marginal value of capital. We can rewrite the HJB as

$$R_iv_i(Z) = Z^\theta + (\mu+\delta)Zv_i'(Z) + \frac{\varepsilon^2}{2}Z^2v_i''(Z) + \lambda_i\left[v_{-i}(Z) - v_i(Z)\right]$$

where $R_i \equiv r_i + \chi + \delta$. If a firm is above it's investment threshold, the marginal value is $v_i(Z) = p^+$and similarly $v_i(Z) = p^-$ below the disinvestment threshold.

We now solve the differential equation in Z each region separately. In the region in which both overlap, we have a system of two second order ODEs. The general solution to the pair of ODEs is

$$v_i(Z) = \eta_iZ^\theta + \sum_{j=1}^4 c_i^jZ^{\gamma_j}$$

with the following definitions

- $H_i(\gamma) \equiv (R_i+\lambda_i) - (\mu+\delta)\gamma - \frac{\varepsilon^2}{2}\gamma(\gamma-1)$

- $\{\gamma_j\}$ are the four roots of the characteristic polynomial $H_1(\gamma) H_2(\gamma) - \lambda_1\lambda_2$
- η_i solve the two equations $H_1(\theta)\eta_1 = 1 + \lambda_1\eta_2$ and $H_2(\theta)\eta_2 = 1 + \lambda_2\eta_1$
- For each j, $\lambda_1 c_2^j = H_1(\gamma_j) c_1^j$. There are thus for unknown constants of integration.

For the region above $\bar{Z}_1$ and below $\bar{Z}_2$ the marginal value satisfy the equations

$$\begin{aligned} v_1(Z) &= p^+ \\ R_2 v_2(Z) &= Z^\theta + (\mu+\delta) Z v_2'(Z) + \frac{\varepsilon^2}{2} Z^2 v_2''(Z) + \lambda_2 \left[p^+ - v_2(Z)\right] \end{aligned}$$

We can solve the ODE for v_2 to get

$$v_2(Z) = \frac{Z^\theta}{H_2(\theta)} + \frac{\lambda_2 p^+}{R_2 + \lambda_2} + \sum_{j=1,2} \psi_{2,j} Z^{\beta_{2,j}}$$

where $\beta_{2,j}$ are the roots of $H_2(\beta) = 0$ and $\psi_{2,j}$ are two more constants of integration.

Lastly for the region above $\underline{Z}_1$ but below $\underline{Z}_2$ the marginal value satisfies the equations

$$\begin{aligned} R_1 v_1(Z) &= Z^\theta + (\mu+\delta) Z v_1'(Z) + \frac{\varepsilon^2}{2} Z^2 v_1''(Z) + \lambda_1 \left[p^- - v_1(Z)\right] \\ v_2(Z) &= p^- \end{aligned}$$

We can solve the ODE for v_1 to get

$$v_1(Z) = \frac{Z^\theta}{H_1(\theta)} + \frac{\lambda_1 p^+}{R_1 + \lambda_1} + \sum_{j=1,2} \psi_{1,j} Z^{\beta_{1,j}}$$

There are now twelve unknowns: four thresholds $\{\bar{Z}_i, \underline{Z}_i\}_{i=1,2}$ and eight constants of integration $\left\{\left\{c_1^j\right\}_{j=1-4}, \{\psi_{i,j}\}_{i,j=1,2}\right\}$. To solve for these unknowns, we appeal to the following conditions:

Smooth Pasting at the upper and lower thresholds:

$$\begin{aligned} v_i(\bar{Z}_i) &= p^+, i = 1,2 \\ v_i(\underline{Z}_i) &= p^-, i = 1,2 \end{aligned}$$

Supercontact at the upper and lower thresholds:

$$\begin{aligned} v_i'\left(\bar{Z}_i\right) &= 0, i = 1, 2 \\ v_i'\left(\underline{Z}_i\right) &= 0, i = 1, 2 \end{aligned}$$

The fact that the value functions are continuously differentiable in the inaction region

$$\begin{aligned} \lim_{Z \to \underline{Z}_2^+} v_1(Z) &= \lim_{Z \to \underline{Z}_2^-} v_1(Z) \\ \lim_{Z \to \underline{Z}_2^+} v_1'(Z) &= \lim_{Z \to \underline{Z}_2^-} v_1'(Z) \end{aligned}$$

$$\begin{aligned} \lim_{Z \to \bar{Z}_1^+} v_2(Z) &= \lim_{Z \to \bar{Z}_1^-} v_2(Z) \\ \lim_{Z \to \bar{Z}_1^+} v_2'(Z) &= \lim_{Z \to \bar{Z}_1^-} v_2'(Z) \end{aligned}$$

With these 12 conditions we can solve for the 12 unknowns.

B.1 Limit as $\lambda_2 \to \infty$

One of the conditions for optimization is

$$v_2\left(\bar{Z}_2\right) - p^+ = 0$$

Note that the for each λ_2, we have

$$(R_2 + \lambda_2)\left(v_2\left(\bar{Z}_2\right) - p^+\right) = 0$$

so that in the limit we have

$$\lim_{\lambda_2 \to \infty} (R_2 + \lambda_2)\left(v_2\left(\bar{Z}_2\right) - p^+\right) = 0 \tag{8}$$

In the middle region ($Z \in \left(\underline{Z}_2, \bar{Z}_1\right)$) we can get at the limit of the value function by looking at the different parts separately. $\left\{\gamma_j\right\}$ are the four roots of the characteristic polynomial $H_1(\gamma) H_2(\gamma) - \lambda_1 \lambda_2$. One can verify that these are

$$\left\{\gamma_j\right\} = -\frac{1}{2}\left(\frac{\mu+\delta}{\varepsilon^2/2} - 1\right) \pm \sqrt{\frac{1}{2}\left(\frac{\mu+\delta}{\varepsilon^2/2} - 1\right) + \frac{1}{\varepsilon^2}\left((R_1 + R_2 + \lambda_1 + \lambda_2) \pm \sqrt{(R_1 + R_2 + \lambda_1 + \lambda_2)^2 - 4\left(R_2\lambda_1 + R_{\cdots}\right.}\right.$$

and also that they satisfy

$$\begin{aligned} j &= 1,2 : H_1\left(\gamma_j\right) = \frac{1}{2}\left[(R_1 - R_2 + \lambda_1 - \lambda_2) + \sqrt{(R_1 + R_2 + \lambda_1 + \lambda_2)^2 - 4\left(R_2\lambda_1 + R_1\lambda_2 + R_1 R_2\right)}\right] \\ j &= 3,4 : H_1\left(\gamma_j\right) = \frac{1}{2}\left[(R_1 - R_2 + \lambda_1 - \lambda_2) - \sqrt{(R_1 + R_2 + \lambda_1 + \lambda_2)^2 - 4\left(R_2\lambda_1 + R_1\lambda_2 + R_1 R_2\right)}\right] \end{aligned}$$

The roots (appropriately scaled) have well defined limits. Order them so that

$$\begin{aligned} \lim_{\lambda_2 \to \infty} \gamma_1 &= s_1 \\ \lim_{\lambda_2 \to \infty} \gamma_2 &= s_2 \\ \lim_{\lambda_2 \to \infty} \frac{\gamma_3}{\sqrt{\lambda_2}} &= -\frac{1}{\sqrt{\varepsilon^2/2}} \\ \lim_{\lambda_2 \to \infty} \frac{\gamma_4}{\sqrt{\lambda_2}} &= \frac{1}{\sqrt{\varepsilon^2/2}} \end{aligned}$$

We now solve for the constants of integration, c_j^i. Recall that for each j, $c_1^j = \frac{H_2(\gamma_j)}{\lambda_2} c_2^j$. Also note that $\lim_{\lambda_2 \to \infty} c_3^j$ and $\lim_{\lambda_2 \to \infty} c_4^j$ must equal zero: if they were finite, the terms $c_i^j Z^{\gamma_j}$ would explode (because $\gamma_j \to \infty$).

I now compute $\frac{H_2(\gamma_j)}{\lambda_2}$. To do this I do the following sequence of calculations. First I compute

the following limit which I will use below

$$
\begin{aligned}
& \lim_{\lambda_2\to\infty} -\lambda_2 + \sqrt{(R_1+R_2+\lambda_1+\lambda_2)^2 - 4\left(R_2\lambda_1 + R_1\lambda_2 + R_1R_2\right)} \\
= & \lim_{\lambda_2\to\infty} \frac{\sqrt{\left(\frac{R_1+R_2+\lambda_1+\lambda_2}{\lambda_2}\right)^2 - \frac{4(R_2\lambda_1+R_1\lambda_2+R_1R_2)}{\lambda_2^2}} - 1}{1/\lambda_2} \\
= & \frac{-2\left(1+\frac{R_1+R_2+\lambda_1}{\lambda_2}\right)\frac{R_1+R_2+\lambda_1}{\lambda_2^2} + 2\frac{4(R_2\lambda_1+R_1R_2)}{\lambda_2^3} + \frac{4R_1}{\lambda_2^2}}{2\sqrt{\left(1+\frac{R_1+R_2+\lambda_1}{\lambda_2}\right)^2 - \frac{4(R_2\lambda_1+R_1\lambda_2+r_1R_2)}{\lambda_2^2}}} / \frac{-1}{\lambda_2^2} \\
= & \frac{2\left(1+\frac{R_1+R_2+\lambda_1}{\lambda_2}\right)(R_1+R_2+R_1) - 2\frac{4(R_2\lambda_1+R_1R_2)}{\lambda_2} - 4R_1}{2} \\
= & -R_1+R_2+\lambda_1
\end{aligned}
$$

We can use this to compute for $j = 1, 2$

$$
\begin{aligned}
\lim_{\lambda_2\to\infty} H_1\left(\gamma_j\right) &= \lim_{\lambda_2\to\infty} \frac{1}{2}\left[(R_1-R_2+\lambda_1-\lambda_2) + \sqrt{(R_1+R_2+\lambda_1+\lambda_2)^2 - 4\left(R_2\lambda_1+R_1\lambda_2+R_1R_2\right)}\right] \\
&= \frac{1}{2}\left[(R_1-R_2+\lambda_1) + (-R_1+R_2+\lambda_1)\right] \\
&= \lambda_1
\end{aligned}
$$

This can be used, in turn to compute the following limit

$$
\lim_{\lambda_2\to\infty} \frac{(R_2+\lambda_2)}{H_2\left(\gamma_j\right)} = \lim_{\lambda_2\to\infty} \frac{R_2+\lambda_2}{\lambda_2}\frac{\lambda_2}{H_2\left(\gamma_j\right)} = \lim_{\lambda_2\to\infty} \frac{R_2+\lambda_2}{\lambda_2}\frac{H_1\left(\gamma_j\right)}{\lambda_1} = 1
$$

$$
\begin{aligned}
\lim_{\lambda_2\to\infty} \lambda_2 - H_2\left(\gamma_j\right) &= \lim_{\lambda_2\to\infty} \lambda_2 - R_2 - \lambda_2 - \left(-\left(\mu+\delta\right)\gamma_j - \frac{\varepsilon^2}{2}\gamma_j\left(\gamma_j-1\right)\right) \\
&= \lim_{\lambda_2\to\infty} R_1 - R_2 + \lambda_1 - R_1 - \lambda_1 - \left(-\left(\mu+\delta\right)\gamma_j - \frac{\varepsilon^2}{2}\gamma_j\left(\gamma_j-1\right)\right) \\
&= \lim_{\lambda_2\to\infty} R_1 - R_2 + \lambda_1 - H_1\left(\gamma_j\right) \\
&= r_1 - r_2
\end{aligned}
$$

And lastly we have

$$
\lim_{\lambda_2\to\infty} (R_2+\lambda_2)\left(\frac{\lambda_2 - H_2\left(\gamma_j\right)}{H_2\left(\gamma_j\right)}\right) = R_1 - R_2
$$

Finally we can use these results to get the following limits

$$\begin{aligned}
\lim_{\lambda_2\to\infty}(R_2+\lambda_2)(\eta_2-\eta_1) &= \lim_{\lambda_2\to\infty}\frac{(R_2+\lambda_2)}{\lambda_2}\frac{R_1-R_2}{H_1(\theta)\frac{H_2(\theta)}{\lambda_2}-\lambda_1} = \frac{R_1-R_2}{\tilde{H}_1(\theta)} \\
\lim_{\lambda_2\to\infty}(R_2+\lambda_2)\left(c_2^j-c_1^j\right) &= \lim_{\lambda_2\to\infty}\frac{(R_2+\lambda_2)}{\lambda_2}\left(\frac{\lambda_2-H_2(\gamma_j)}{H_2(\gamma_j)}\right)c_1^j = (R_1-R_2)\,c_1^j
\end{aligned}$$

where $\tilde{H}_1(\gamma) = R_1-(\mu+\delta)\gamma-\frac{\varepsilon^2}{2}\gamma(\gamma-1)$. We then have

$$\lim_{\lambda_2\to\infty}(R_2+\lambda_2)(v_2(Z)-v_1(Z)) = (R_1-R_2)\,v_1(Z)$$

Since this is true for all $Z\in\left(\underline{Z}_2,\bar{Z}_1\right)$, we have

$$\lim_{\lambda_2\to\infty}\lim_{Z\to\bar{Z}_1^-}(R_2+\lambda_2)\left(v_2(Z)-p^+\right) = (r_1-r_2)\,p^+ < 0$$

Combining this with (8) implies that $\bar{Z}_2 > \bar{Z}_1$.

C Dispersion and Jensen's Inequality

Relative TFP can be written as $\frac{TFP_s}{A_s} = \left(\sum_i B_i e^{(\sigma_s-1)(\tau_s-\tau_i)}\right)^{\frac{1}{\sigma_s-1}}$. If factor shares are defined to satisfy $\alpha_{Ks} = \sum \alpha_{Ki}\frac{P_iY_i}{P_sY_s}$ than $TFP_s \le A_s$. In the course of the proof, I will make use of the fact that the revenue shares can be written as

$$\frac{P_iY_i}{P_sY_s} = \left(\frac{P_i}{P_s}\right)^{1-\sigma_s} = \frac{(\lambda_iA_i)^{\sigma_s-1}\,e^{(1-\sigma_s)\tau_i}}{\sum_i(\lambda_iA_i)^{\sigma_s-1}\,e^{(1-\sigma_s)\tau_i}} = \frac{B_ie^{-(\sigma_s-1)\tau_i}}{\sum_i B_ie^{-(\sigma_s-1)\tau_i}} \tag{9}$$

and similarly $\frac{P_sY_s}{PY} = \frac{B_se^{-(\gamma-1)\tau_s}}{\sum_i B_se^{-(\gamma-1)\tau_s}}$

The first order condition with respect to capital from the sector level problem

$$
\begin{aligned}
e^{-\tau_{Ks}} &= \frac{\sigma_s}{\sigma_s - 1}\frac{1}{\alpha_{Ks}}\frac{RK_s}{P_sY_s} \\
&= \frac{\sigma_s}{\sigma_s - 1}\frac{1}{\alpha_{Ks}}\sum_i \frac{RK_i}{P_iY_i}\frac{P_iY_i}{P_sY_s} \\
&= \frac{1}{\alpha_{Ks}}\sum_i \alpha_{Ki}e^{-\tau_{Ki}}\frac{P_iY_i}{P_sY_s}
\end{aligned}
$$

where the last equality use the first order condition from the firm's problem and the expression for revenue share (9). Let $\omega_{Ki} \equiv \frac{\alpha_{Ki}P_iY_i}{\alpha_{Ks}P_sY_s}$.

$$
e^{-\tau_{Ks}} = \sum_i e^{-\tau_{Ki}}\omega_{Ki}
$$

Under the assumption that $\alpha_K = \sum \alpha_{Ki}\frac{P_iY_i}{P_sY_s}$, we have $\sum \omega_i = 1$. Since e^x is convex, Jensen's inequality gives $E\left[e^x\right] \geq e^{E(x)}$, so we can write

$$
e^{-\tau_{Ks}} \geq \exp\left[\sum_i \left[-\tau_{Ki}\right]\omega_{Ki}\right]
$$

Lastly we can write

$$
e^{(\sigma_s-1)\alpha_{Ks}\tau_{Ks}} \leq \exp\left[\alpha_{Ks}\left(\sigma_s - 1\right)\sum_i \tau_{Ki}\omega_{Ki}\right]
$$

We can derive similar expressions for labor and for materials. With these we can write

$$
\begin{aligned}
e^{(\sigma_s-1)\tau_s} &= e^{(\sigma_s-1)(\alpha_{Ks}\tau_{Ks}+\alpha_{Ls}\tau_{Ls}+\alpha_{Ms}\tau_{Ks})} \\
&\leq \exp\left[\left(\sigma_s - 1\right)\sum_i \alpha_{Ks}\tau_{Ki}\omega_{Ki} + \alpha_{Ls}\tau_{Li}\omega_{Li} + \alpha_{Ms}\tau_{Mi}\omega_{Mi}\right]
\end{aligned}
$$

We can substitute back in the expressions for ω_{Ki}, ω_{Li}, and ω_{Mi}, to get in to get

$$
\begin{aligned}
e^{(\sigma_s-1)\tau_s} &\leq \exp\left[\left(\sigma_s - 1\right)\sum_i \left(\alpha_{Ki}\tau_{Ki} + \alpha_{Li}\tau_{Li} + \alpha_{Mi}\tau_{Mi}\right)\frac{P_iY_i}{P_sY_s}\right] \\
&= \exp\left[\sum_i \left(\sigma_s - 1\right)\tau_i\frac{P_iY_i}{P_sY_s}\right]
\end{aligned}
$$

Using Jensen's inequality one more time, we have

$$e^{(\sigma_s-1)\tau_s} \leq \sum_i \left[e^{(\sigma_s-1)\tau_i} \frac{P_i Y_i}{P_s Y_s} \right]$$

Lastly, using the expression for the revenue share (9), we get

$$e^{(\sigma_s-1)\tau_s} \leq \sum_i \left[e^{(\sigma_s-1)\tau_i} \frac{B_i e^{(1-\sigma_s)\tau_i}}{\sum_i B_i e^{(1-\sigma_s)\tau_i}} \right] = \sum_i \left[\frac{B_i}{\sum_i B_i e^{(1-\sigma_s)\tau_i}} \right] = \frac{1}{\sum_i B_i e^{(1-\sigma_s)\tau_i}}$$

This implies

$$\sum_i B_i e^{(\sigma_s-1)(\tau_s-\tau_i)} \leq 1$$

Aggregate Level

At the aggregate level we want to show

$$\left(\sum B_s \left(\frac{TFP_s}{A_s} \right)^{\gamma-1} e^{(\gamma-1)(\tau-\tau_s)} \right)^{\frac{1}{\gamma-1}} \leq \left(\sum_s B_s \left(\frac{TFP_s}{A_s} \right)^{\gamma-1} \right)^{\frac{1}{\gamma-1}}$$

This can be done if we define $\log \mu = \sum_s \log \mu_s \frac{P_s Y_s}{PY}$ where $\mu_s = \frac{\sigma_s}{\sigma_s - 1}$ and $\mu = \frac{\sigma}{\sigma-1}$. The capital wedge for the aggregate level can be written as

$$\begin{aligned} e^{-\tau_K} &= \frac{\sigma}{\sigma-1} \frac{1}{\alpha_K} \frac{RK}{PY} \\ &= \frac{\sigma}{\sigma-1} \frac{1}{\alpha_K} \sum_i \frac{RK_s}{P_s Y_s} \frac{P_s Y_s}{PY} \\ &= \frac{1}{\alpha_K} \sum_i \alpha_{Ks} e^{-\tau_{Ks} + \log \frac{\mu}{\mu_s}} \frac{P_s Y_s}{PY} \\ &= \sum_i e^{-\tau_{Ks} + \log \frac{\mu}{\mu_s}} \omega_{Ks} \end{aligned}$$

where $\omega_{Ks} = \frac{\alpha_{Ks} P_s Y_s}{\alpha_K PY}$. With the assumptions that $\alpha_K = \sum_s \alpha_{Ks} \frac{P_s Y_s}{PY}$, we have $\sum_s \omega_{Ks} = 1$.

$$e^{-\tau_K} \geq \exp \left[\sum_s -\tau_{Ks} + \log \left(\frac{\mu}{\mu_s} \right) \omega_{Ks} \right]$$

Lastly we can write

$$e^{(\gamma-1)\alpha_K\tau_{Ks}} \leq \exp\left[\alpha_K(\gamma-1)\sum_s\left(\tau_{Ks}+\log\left(\frac{\mu_s}{\mu}\right)\right)\omega_{Ks}\right]$$

We can derive similar expressions for labor and for materials. With these we can write

$$\begin{aligned} e^{(\gamma-1)\tau} &= e^{(\gamma-1)(\alpha_K\tau_K+\alpha_L\tau_L+\alpha_M\tau_K)} \\ &\leq \exp\left[(\gamma-1)\sum_s \alpha_K\tau_{Ks}\omega_{Ks}+\alpha_L\tau_{Ls}\omega_{Ls}+\alpha_M\tau_{Ms}\omega_{Ms}+\log\left(\frac{\mu_s}{\mu}\right)(\alpha_K\omega_{Ks}+\alpha_L\omega_{Ls}+\alpha_M\omega_{Ms})\right] \end{aligned}$$

We can substitute back in the expressions for ω_{Ks}, ω_{Ls}, and ω_{Ms}, to get in to get

$$\begin{aligned} e^{(\gamma-1)\tau} &\leq \exp\left[(\gamma-1)\sum_s(\alpha_{Ks}\tau_{Ks}+\alpha_L s\tau_{Ls}+\alpha_{Ms}\tau_{Ms})\frac{P_sY_s}{PY}+\log\left(\frac{\mu_s}{\mu}\right)(\alpha_{Ks}+\alpha_{Ls}+\alpha_{Ms})\frac{P_sY_s}{PY}\right] \\ &= \exp\left[(\gamma-1)\sum_s\left(\tau_s+\log\left(\frac{\mu_s}{\mu}\right)\right)\frac{P_sY_s}{PY}\right] \end{aligned}$$

Also, under the assumption that $\log\mu = \sum_s \log\mu_s\frac{P_sY_s}{PY}$, this becomes

$$e^{(\gamma-1)\tau} \leq \exp\left[\sum_s(\gamma-1)\tau_s\frac{P_sY_s}{PY}\right]$$

Using Jensen's inequality one more time, we have

$$e^{(\gamma-1)\tau_s} \leq \sum_s\left[e^{(\gamma-1)\tau_s}\frac{P_sY_s}{PY}\right]$$

Using $\frac{P_sY_s}{PY} = \frac{X_se^{-(\gamma-1)\tau_s}}{\sum X_se^{-(\gamma-1)\tau_s}}$

$$\begin{aligned} e^{(\gamma-1)\tau_s} &\leq \sum_s\left[e^{(\gamma-1)\tau_s}\frac{X_se^{-(\gamma-1)\tau_s}}{\sum X_se^{-(\gamma-1)\tau_s}}\right] \\ &= \frac{1}{\sum X_se^{-(\gamma-1)\tau_s}} \end{aligned}$$

Rearranging gives

$$\left(\sum X_se^{(\gamma-1)(\tau-\tau_s)}\right)^{\frac{1}{\gamma-1}} \leq 1$$

$$X_s = \frac{B_s \left(\frac{TFP_s}{A_s}\right)^{\gamma-1}}{\sum_s B_s \left(\frac{TFP_s}{A_s}\right)^{\gamma-1}}$$

$$\left(\sum B_s \left(\frac{TFP_s}{A_s}\right)^{\gamma-1} e^{(\gamma-1)(\tau-\tau_s)}\right)^{\frac{1}{\gamma-1}} \leq \left(\sum_s B_s \left(\frac{TFP_s}{A_s}\right)^{\gamma-1}\right)^{\frac{1}{\gamma-1}}$$