ECONOMIC STUDIES

**DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
206**

_____

**Dynamic Investment Models, Employment Generation and Productivity
– Evidence from Swedish Data**

Yoshihiro Sato

UNIVERSITY OF GOTHENBURG

To my dear parents, grandparents, sister and brother.

Contents

Acknowledgements

My journey as a PhD student started in September 2004. Since the beginning, it has been full of sweat, tears, and occasional despair, but also hope.

First of all, I would like to thank my supervisor Måns Söderbom for his engagement, advice, assistance, encouragement, and a lot more. During the work with this thesis, I have encountered many hurdles and have sometimes been lost in the jungle of econometric models and statistical problems. However, every time I met with Måns to get help, he was patient, positive and constructive, and always helped me work through the problems. Måns, your guidance has taught me a lot. I have been truly fortunate to have you as my supervisor.

I am very grateful to Arne Bigsten, who connected me to Måns. I wish also thank to Johan Stennek for his support and encouragement.

I also received a lot help and support from Lennart Hjalmarsson and Roger Wahlberg. I wish they were here with us and I'd be able to thank them.

Also, thanks to my classmates Pelle Ahlerup, Annika Lindskog, Ann-Sofie Isaksson, Miyase Yesim Köksal-Ayhan, and Niklas Jakobsson. We started the PhD program together and have had a good time.

I also wish to thank Matilda Orth, Florin Maican, Eyerusalem Siba and Daniela Roughsedge for support and advice on the topic of productivity analysis. To Dario Pozzoli, thanks for your help with STATA programming and serving as opponent at my final seminar.

Among other colleagues at the department, I would especially like to thank Sven Tengstam, Heather Congdon Fors, Anders Boman, Måns Nerman, Björn Sund, Gustav Hansson, Anna Widerberg, Håkan Eggert, Katarina Nordblom, Åsa Löfgren, Karin Backteman, Selma Oliveira, Katarina Renström and not least Elizabeth Földi for personal support and nice conversations at the lunch table. I have had many insightful conversations with Johan Lönnroth and Bo Sandelin. I am very glad that Johan visited me in Japan. Hans Bjurek, Margareta Westberg, Per-Åke Andersson, Hans Hansson, Lars-Göran Larsson, Evert Köstner, Wlodek Bursztyn, and Mattias Sundén have helped me with my teaching of undergraduates on several occasions. Many other colleagues have supported me directly and indirectly, as well.

I am also grateful to Eva-Lena Neth-Johansson, Jeanette Saldjoughi, Carita Olausson and Mona Jönefors for administrative assistance. Debbie Axlid has been an excellent editor of my five papers. Thank you.

Saving the best for last, I wish to send a great big thank you to my dear parents, grandparents, sister and brother. Thank you all for showing understanding for my keen desire for studying, researching, and completing a PhD degree in Sweden – a country far away from home. Grandmother, it makes me smile every time I remember that you once sent me a postcard with a message starting "How's life in Switzerland…"

This thesis would not have seen the light of the day without the help I have received from so many kind and brilliant people.

Yoshihiro Sato

Gothenburg, September 19, 2012

Summary of the thesis

The thesis consists of five self-contained papers. Two of them are co-authored with Måns Söderbom.

Paper 1: *Effects of shocks and uncertainty on capital and labor in a real options model with variable capacity utilization*

This study analyzes the effects of uncertainty on investment, capital use and labor demand, in both the short run and the long run, in dynamic capital models where firms have incentives not to operate at full capacity. We have modified the real options model with full irreversibility to incorporate variable capacity utilization. After obtaining a numerical solution to the investment function, we analyze the difference in the prediction between the real options model with full capacity and our models. In the short run, the firm is less cautious for new investment and accumulates more capital than in a real options model with full capacity utilization. In the long run, the firm tends to build a larger capital stock while capital use declines as uncertainty increases. This suggests inefficiency in the form of excessive capital accumulation. Implications for labor demand are also discussed.

Paper 2: *Capital adjustment under variable capacity utilization*
(co-authored with Måns Söderbom)

This study generalizes the existing real options model so as to accommodate fluctuations in capacity utilization, and then analyzes the effects of uncertainty in this modified model. Using both simulated data and Swedish 2-digit sectoral data for the manufacturing industry, we show that the main argument of the real options model – that the responses of the capital stock to demand shocks are weaker at higher levels of uncertainty – still holds, implying that policy stimuli have limited effects on investment in the immediate aftermath of an uncertainty shock. On the other hand, we find that actual capital use (active capital) flexibly responds to demand shocks through adjustments of capacity utilization even at high levels of uncertainty, which suggests that policy stimuli have positive impacts on production activities even at high uncertainty.

Paper 3: *Are larger firms more productive because of scale economies? – Evidence from Swedish register data*
(co-authored with Måns Söderbom)

This study investigates the factors driving higher labor productivity for large firms, using Swedish register-based microdata for the mining and manufacturing industries covering more than 28,000 firms during 1997-2006. We estimate translog production functions using dynamic panel approaches and the approach proposed by Ackerberg, Caves, and Frazer. The results show that micro and small firms operate under (locally) increasing returns to scale while medium and large firms face decreasing returns to scale. Scale elasticity decreases from 1.15 to 0.97, suggesting that scale effects are not the answer to our question. Further investigation shows that production technology is approximated by a non-homothetic function and that larger firms operate with more capital-intensive technology while the factor price ratio is constant, which drives the productivity difference in favor of larger firms.

Paper 4: *Employment generation and productivity contribution of entrepreneurial firms compared to large incumbents*

Previous studies have reported that young and small so-called entrepreneurial firms have disproportionately large impacts on both employment generation and productivity growth. However, these positive impacts are conditional on the firms' survival. Many studies show that young and small firms have high mortality. In our study, we investigate the contributions of entrepreneurial firms to employment generation and productivity growth after taking the high mortality into account. We find that, although young and small firms are less likely to survive, they contribute more than other firms to both employment generation and aggregate productivity, simply because the survivors perform eminently.

The last paper, **Paper 5**, is a short complement to **Paper 4**.

Paper 5: *Initial firm size and post-entry growth in size and productivity*

Previous studies show that smaller entrants exhibit higher growth rates in terms of size (number of employees). Using register-based firm-level data for the Swedish mining and manufacturing industries, this study compares the development in size and productivity between a group of firms that started their business in 1998 and a group of firms that had been in business for at least 10 years in 1998. The results show that there is also a similar negative relationship between firms' initial size and post-entry productivity development. The average total factor productivity of the entrants is initially 15 percent lower than that of the incumbents and the difference becomes insignificant after three years. Regarding the growth pattern conditional on initial size, the entrants with one initial employee caught up with the incumbents of similar initial size already in the second year, and gained a lead in the fifth year. It takes three years for the productivity gap between the entrants with 10 initial employees and the incumbents with the same size to disappear. The entrants with more than 20 initial employees never caught up with the entrants of the same size during the nine years analyzed in this study.

Paper 1

Effects of Shocks and Uncertainty on Capital and Labor in a Real Options Model with Variable Capacity Utilization

Yoshihiro Sato
University of Gothenburg

Abstract

This study analyzes the effects of uncertainty on investment, capital use and labor demand, in both the short run and the long run, in dynamic capital models where firms have incentives not to operate at full capacity. We have modified the real options model with full irreversibility to incorporate variable capacity utilization. After obtaining a numerical solution to the investment function, we analyze the difference in the prediction between the real options model with full capacity and our models. In the short run, the firm is less cautious for new investment and accumulates more capital than in a real options model with full capacity utilization. In the long run, the firm tends to build a larger capital stock while capital use declines as uncertainty increases. This suggests an inefficiency in the form of excessive capital accumulation. Implications for labor demand are also discussed.

1 Introduction

Recent developments in the theory on dynamic investments under uncertainty highlight the impact of a completely or partly irreversible investment.[1] The theory predicts that firms are more reluctant to invest since there exists an option value to delay an investment decision in order to wait for the arrival of new information. Consequently, firms have a cost "threshold" for new investments that is higher than the Jorgensonian user cost of capital, which leads to less accumulation of capital in the short run. Abel and Eberly (1999) call it

---

[1]See, e.g., Abel (1983), Pindyck (1988), Pindyck (1993), Bertola and Caballero (1994), Dixit and Pindyck (1994), Abel and Eberly (1996, 1999), Leahy and Whited (1996), Guiso and Parigi (1999), Abel (2002), Bloom et al.(2003, 2007), Bond and Lombardi (2006), Cooper and Haltiwanger (2006), Bloom (2009), Chirinko and Schaller (2009), and Bond et al.(2011).

the "user cost" effect.[2] Previous studies show that the difference between the Jorgensonian user cost and the threshold, and hence the cautiousness, is larger as the degree of uncertainty increases.

In models of investment under uncertainty, net investment is the only source of variation in capital input in a firm. The firm tries to anticipate future demand shocks and makes an investment plan to maximize the expected profits. However, due to a time lag before the installed capital becomes available as well as uncertainty of future demand shocks, there is always a gap between the *ex post* desired capital stock and the current capital stock. This is the reason why the firm is cautious in undertaking new investment.

In reality, however, we usually observe considerable fluctuation in capacity utilization over the business cycle. For instance, the Swedish manufacturing industry exhibits a capacity utilization rate that has fluctuated between 80 and 92 percent during the last two decades. Note also that the capacity utilization rate never reached 100 percent. These facts suggest that the capital stock is rather flexible internally, in the sense that the firm is able to adjust its capital input to the current demand shock by varying capacity utilization. This mechanism is typically ignored in modern models of investment.

In this study, we analyze the effects of uncertainty on capital and labor in both the short run and the long run in a model where firms have incentives not to operate at full capacity.[3] The degree of capacity utilization as well as the amount of new investment are therefore decision variables for the firms in this model. We compare the result with the predictions by a dynamic investment model proposed by Abel and Eberly (1999) (henceforth the AE model).

Allowing for variable capacity utilization in dynamic capital theory will have a direct impact on the capital accumulation behavior of the firm. Firstly, because the "user cost" effect becomes lower than predicted by the previous real options model, reluctance toward new investments will now be less. Secondly, it also affects labor costs. Since reduction of capital use in a recession involves reduction of labor input as well, a forward-looking firm will see lower binding costs related to new investment. Consequently, a firm in our model will respond more with investment to a positive demand shock. The difference between reluctance in our model and in the AE model will be smaller as the degree of

---

[2] This is because the wedge between the threshold and the traditional Jorgensonian user cost is derived from the uncertainty-revised user cost of capital that incorporates the effect of the irreversibility constraints.

[3] In previous models, capital is assumed to be free to use once it has been installed. Therefore, there is no reason to refrain from making full use of it even in a recession.

uncertainty increases.

To let capacity utilization vary over the business cycle, we propose two different model settings. Our first model is built upon a Cobb-Douglas production function. We assume that there is a cost related to capital use, in addition to the purchase cost. We call it a "maintenance cost". The firm needs to pay this cost proportionally to capital use. This is analogous to an assumption that the capital depreciation rate is a positive function of capital use. Under this assumption, a firm does not always operate at full capacity simply because it is costly to do so when facing low demand.

Our second model requires no such costs. We instead assume a Leontief production function. Labor and capital are combined at a fixed ratio. Both production factors can be used less when the firm faces a negative demand shock. This is because labor, which is assumed to be flexible, is costly to use and the ratio between capital input and labor input is fixed. As a result, part of the capital stock can be redundant, which is expressed by variations in the capacity utilization rate.

We then analyze both the short- and long-run behavior of the firm through numerical simulation based on each of our models. The AE model serves as a benchmark. Abel and Eberly (1999) construct this dynamic investment model to theoretically determine the sign of the effect of irreversibility and uncertainty on expected long-run capital stock. They point out that both the user cost effect and a "hangover" effect operate under irreversibility constraints. The latter refers to the effect of irreversibility as irreversibility prevents the firm from getting rid of capital when its marginal revenue product is low. As a consequence of these two effects, Abel and Eberly prove that the sign of the effect of irreversibility and uncertainty on expected long-run capital stock is ambiguous: at higher uncertainty, the user cost effect dominates and the firm accumulates less capital than in the model without irreversibility. Although the AE model is used by Abel and Eberly for theoretical analysis of the long-term effects, we use it to analyze the short-term behavior of firms, using numerical simulation.

Our results show that, in the short run, the firm is less cautious with respect to new investments and accumulates more capital than in the AE model with full capacity utilization. On the other hand, the firm reduces active capital, and thus capacity utilization, in a recession. The reduction of active capital implies that labor input is also reduced. In contrast, active capital in the AE model with full utilization stays high even in a recession, and labor input thereby stays

the same or decreases marginally.

In the long run, our results exhibit that the firm with variable capacity utilization tends to build a larger capital stock as uncertainty increases. This is in contrast to what the AE model predicts. As a result, the hangover effect dominates. This result is explained by two factors. The first one is only applicable to the maintenance cost model. When uncertainty increases, the expected rate of capacity utilization decreases, which leads to a lower effective user cost of capital, and hence more capital accumulation.

The other factor is applicable to both the maintenance cost model and the Leontief model, and deals with labor costs. The reluctance toward new investment is not only due to the capital costs that are sunk, but also to the labor costs that are related to capital use. Since the firm is able to reduce active capital and thereby labor input, the total binding costs of new investments, including the user cost of capital and the labor cost, are lower. Moreover, the expected value decreases when the firm is more likely to face recessions at higher uncertainty.

Another result in our analysis is that the firm's active capital decreases as uncertainty increases, though capital accumulation at the same time intensifies, as we discussed above. This implies that the expected capacity utilization is a decreasing function of uncertainty.

This study is structured as follows. In Section 2, we review previous studies. In Section 3, we outline the maintenance cost model and the Leontief model we propose, as well as the AE model as a benchmark. In Section 4, we present numerical estimation of the investment function according to the different model assumptions, and discuss the short-run and long-run impacts of variable capacity utilization and uncertainty on investment, capital use, and labor demand. Section 5 concludes the study.

2   Previous studies

To our knowledge, relatively few previous studies in capital investment theory incorporate variable capacity utilization. Taubman and Wilkinson (1970) analyze a firm's behavior in terms of capital use and investment through a cost minimization problem. The depreciation rate of capital is an increasing function of capital use[4], and the user cost of capital is thereby increasing in the intensity

---

[4]In previous studies, the term "capital utilization" is often used synonymously with "capacity utilization" when only capital goods are assumed to be quasi-fixed.

of capital use. Their analysis is static without any dynamic aspect.

Smith (1969, 1970) outlines a similar neoclassical model with the depreciation rate depending on capital utilization, but also introduces uncertainty. In his model, firms make decisions in two steps. First they choose the capital stock level before current demand is known. Then, once current demand is known, they choose the input level of flexible factors and the capital utilization rate. His focus is on the long-run effect in the equilibrium and he finds a negative relationship between uncertainty and the expected capital utilization rate. He also finds a positive relationship between uncertainty and the optimal capital stock if the production function is a Cobb-Douglas.

Abel (1981) develops a more extensive dynamic investment model, where he defines capital as well as labor as quasi-fixed factors, and allows utilization of both factors to vary jointly. As the labor cost depends on labor utilization, which is related to capital utilization, instantaneous maximization of the profit function brings variation in capital utilization. His analysis concerns the firm's behavior in the short run, and he finds that the response of capital investment to a demand shock is positively related to capacity utilization.

Within the literature on dynamic interrelated factor demand models, Bischoff and Kokkelenberg (1987) construct models where, similar to Taubman and Wilkinson (1970) and Smith (1969, 1970), the depreciation rate depends on the intensity of capital use and the firm thereby actively chooses the optimal level of capital utilization. Morrison (1986), Bernstein and Nadiri (1988), Prucha and Nadiri (1996), and Fousekis and Stefanou (1996) argue that the capacity utilization rate can be measured by divergence between the long-run and short-run equilibria.

Another field that takes capacity utilization into account is dynamic stochastic general equilibrium models. Christiano et al.(2005) incorporate variable capital utilization into a dynamic growth model of general equilibrium theory. A problem with the preceding models in this field is a large rise in marginal costs and prices after a positive monetary policy shock. They find that allowing for variation in the services of capital for a given capital stock reduces the rise in marginal costs, and thus better replicates inflation inertia and output persistence under the assumption of only moderate wage and price stickiness.

Regarding the two model settings used in the present study, the maintenance cost model is similar to the models by Taubman and Wilkinson (1970) and Smith (1969, 1970). The maintenance costs that the firm needs to pay for used capital can be interpreted as the costs to replace depreciated capital. Our

second model with a Leontief production function is closely related to the model by Abel (1981), in that capital input and labor input are utilized at a fixed ratio. However, we will not analyze the effect of the time lag between an investment decision and realization of actual demand as in Smith (1969, 1970). We instead assume that purchased capital goods are available for production already in the same period.

3 A dynamic investment model under uncertainty

Abel and Eberly (1999) outline a simple dynamic model of a representative firm with Cobb-Douglas production technology facing an isoelastic demand curve. They compare capital accumulation in the long run when a perfect irreversibility constraint is imposed and when it is not.

We briefly introduce a discrete approximation of their continuous-time model (the AE model) with and without an irreversibility constraint. We then two different models that make capacity utilization vary over the business cycle. Our first model adds a maintenance cost of capital to the AE model. In our second model, the firm has a Leontief production technology that combines capital and labor inputs at a fixed ratio.

3.1 The Abel-Eberly (AE) model

On the demand side, we assume a firm that faces an isoelastic demand curve

$$Q_t = X_t P_t^{-\epsilon}, \tag{1}$$

where Q_t is the quantity of output demanded, P_t is the price of output, and $\epsilon > 1$ is the price elasticity of demand. X_t is a stochastic demand shock, which is the only source of uncertainty in the model. This is a "horizontal demand shock" in that it shifts the quantity demanded Q_t for any given output price P_t. We assume that X_t evolves according to a geometric Brownian motion with mean μt and variance $\sigma^2 t$. According to Ito's lemma, the log of X_t follows a Brownian motion with mean $\left(\mu - 0.5\sigma^2\right) t$ and variance $\sigma^2 t$:

$$\ln X_t \equiv x_t = x_{t-1} + \tilde{\mu} + u_t, \tag{2}$$

where $\tilde{\mu} = \mu - 0.5\sigma^2$, $u_t = iid.N\left(0, \sigma^2\right)$, and $x_0 = 0$. The variance of the demand shock σ^2 indicates the level of uncertainty that the firm faces.

On the supply side, we assume that a representative firm produces output Q_t using the Cobb-Douglas production technology with constant returns to scale

$$Q_t = L_t^{1-\alpha} K_t^{\alpha}, \tag{3}$$

where L_t is labor, K_t is the capital stock employed in the production process, and α is the cost share of capital. Labor is assumed to be flexible so that the firms chooses L_t to maximize its operating profit $\pi_t = P_t Q_t - w L_t$, where w is the wage rate, which is assumed to be constant. The first-order condition leads to optimal labor as

$$L^*\left(K_t, X_t\right) = w^{-\gamma\epsilon} \left(\frac{\gamma\epsilon - 1}{\gamma\epsilon}\right)^{\gamma\epsilon} X_t^{\gamma} K_t^{1-\gamma} \tag{4}$$

and the operating profit becomes

$$
\begin{aligned}
\pi^*\left(K_t, X_t\right) &= w^{-\gamma\epsilon} \left(\frac{\gamma\epsilon - 1}{\gamma\epsilon}\right)^{\gamma\epsilon} \frac{w}{\gamma\epsilon - 1} X_t^{\gamma} K_t^{1-\gamma} \\
&= h X_t^{\gamma} K_t^{1-\gamma}, \tag{5}
\end{aligned}
$$

where

$$\gamma = \frac{1}{1 + \alpha\left(\epsilon - 1\right)} \tag{6}$$

and

$$h = w^{-\gamma\epsilon} \left(\frac{\gamma\epsilon - 1}{\gamma\epsilon}\right)^{\gamma\epsilon} \frac{w}{\gamma\epsilon - 1}. \tag{7}$$

The capital stock equals the accumulation of past and current investment with depreciation at the fixed rate δ. We assume that current investment is immediately added to the capital stock without any delay, since our study is based upon a discrete approximation of Abel and Eberly's continuous-time model. Accordingly, the capital stock evolves as

$$K_t = \left(1 - \delta\right) K_{t-1} + I_t. \tag{8}$$

In a frictionless case, where we ignore irreversibility constraints, the firm is able to adjust its capital stock immediately both upwards and downwards after observing realization of X_t. Uncertainty plays no role in the investment

decision. The dynamic optimization problem is

$$V\left(K_{t-1}, X_t\right) = \max_{I_t} hX_t^\gamma K_t^{1-\gamma} - qI_t + \beta E_t\left[V\left(K_t, X_{t+1}\right)\right], \qquad (9)$$

where q is the purchase cost of capital. The solution to the problem constitutes the frictionless optimal level of the capital stock K_t^{f*}, which is expressed by

$$K_t^{f*} = \left\{\frac{h\left(1-\gamma\right)}{q\left[1-\beta\left(1-\delta\right)\right]}\right\}^{\frac{1}{\gamma}} X_t. \qquad (10)$$

The user cost of capital, which is known as the Jorgesonian frictionless user cost, is

$$c = q\left[1 - \beta\left(1-\delta\right)\right]. \qquad (11)$$

When adding an irreversibility constraint, the dynamic optimization problem for the firm is

$$V\left(K_{t-1}, X_t\right) = \max_{I_t \geq 0} hX_t^\gamma K_t^{1-\gamma} - qI_t + \beta E_t\left[V\left(K_t, X_{t+1}\right)\right]. \qquad (12)$$

The solution to the dynamic optimization problem is expressed by

$$I_t = I\left(K_{t-1}, X_t; \sigma\right). \qquad (13)$$

In this model, optimal investment reflects two mechanisms: the user cost and the hangover effects. Since the firm is now unable to sell capital stock, it is constrained by its own investment decision in the past. This may result in a capital stock level higher than what is optimal in the frictionless model, and is called a hangover effect. The other side of the same coin is a user cost effect. The firm knows that it is unable to cut back on the capital stock in an economic downturn. Because of its cautious decisions regarding investments, the responsiveness of investments to a positive demand shock is weaker. As already mentioned, Abel and Eberly prove that the sign of the effect of irreversibility and uncertainty on expected long-run capital stock is ambiguous as the consequence because these two effects counteract each other.

3.2 Model with varying capacity utilization (maintenance cost model)

We now let active capital K_t^a depart from the installed capital stock K_t. We introduce the maintenance cost, m, in addition to the purchase cost. The dynamic optimization problem for the firm, then, is

$$V\left(K_{t-1}, X_t\right) = \max_{\substack{I_t \geq 0 \\ K_t^a \leq K_t}} hX_t^\gamma \left(K_t^a\right)^{1-\gamma} - mK_t^a - qI_t + \beta E_t \left[V\left(K_t, X_{t+1}\right)\right], \quad (14)$$

where q is the purchase cost of capital. The maintenance cost is the cost using the capital. It can also be interpreted as the cost to replace deteriorated capital to keep the capital stock at the same level. An extreme case where $m = q$ is equivalent to replacing all of the used capital. In other words, the depreciation rate is 100 percent. When m is instead one-fifth of q (i.e., $m = 0.20q$), it can be interpreted as that the firm needs to replace one-fifth of the used capital. The depreciation rate is 20 percent. Thus, the ratio m/q is equivalent to the depreciation rate.[5]

The optimal input of labor is

$$L^*\left(K_{t-1}, X_t\right) = w^{-\gamma\epsilon} \left(\frac{\gamma\epsilon - 1}{\gamma\epsilon}\right)^{\gamma\epsilon} X_t^\gamma \left(K_t^a\right)^{1-\gamma}. \quad (15)$$

Optimal labor input thus depends on active capital and not the capital stock as in Eq.(4) of the AE model.

Choice of K_t^a is derived by static optimization of the operating profit function as

$$\max_{K_t^a \leq K_t} P_t Q_t - wL_t^* - mK_t^a \quad (16)$$

$$= \max_{K_t^a \leq K_t} hX_t^\gamma \left(K_t^a\right)^{1-\gamma} - mK_t^a. \quad (17)$$

The solution is

$$K_t^{a*} = \min\left(K_t, \ \tilde{K}_t\right), \quad (18)$$

where

$$\tilde{K}_t = \left[\frac{h\left(1-\gamma\right)}{m}\right]^{\frac{1}{\gamma}} X_t. \quad (19)$$

---

[5]We explore some more extreme cases. When $m/q = \infty$, it means that the firm can freely rent capital, which is equivalent to the frictionless model. When $m/q = 0$, it is equivalent to the AE model with total irreversibility.

The firm's dynamic optimization problem, Eq.(14), can be rewritten as

$$V\left(K_{t-1}, X_t\right) = \max_{I_t \geq 0} h X_t^\gamma \left(K_t^{a*}\right)^{1-\gamma} - m K_t^{a*} - q I_t + \beta E_t\left[V\left(K_t, X_{t+1}\right)\right]. \quad (20)$$

The solution is expressed by

$$I_t = I\left(K_{t-1}, X_t; \sigma\right). \quad (21)$$

The Jorgesonian frictionless user cost is

$$c = q\left(\frac{1}{1-\delta} - \beta\right) + m. \quad (22)$$

Note that the capital user cost is higher by m compared to Eq. (11). The utilization rate U_t is expressed by

$$U_t \equiv \frac{K_t^{a*}}{K_t} \leq 1. \quad (23)$$

Note that K_t is a control variable through I_t because $K_t = (1-\delta)K_{t-1} + I_t$. The firm is willing to invest whenever $\tilde{K}_t$ exceeds K_t, since a larger K_t^{a*} means a higher operating profit π. However, since new investment incurs the purchase cost q per unit and investment cannot be reversed, the firm will not always fill the whole gap between $\tilde{K}_t$ and K_t.

Note also that the firm makes new investments only when $U_t = 1$. This is because the firm can expand the capital stock immediately whenever it is necessary and there is thus no need for the firm to anticipate a future positive demand shock when $U_t < 1$.

We now show that the variation in capacity utilization U_t is negatively related to the ratio $\tilde{K}_t/K_t^*$ and positively related to the ratio m/q. In the frictionless case where there is no irreversibility constraint, the firm will keep the frictionless optimal capital level that is expressed by

$$K_t^{f*} = \left\{\frac{h\left(1-\gamma\right)}{m + q\left[1 - \beta\left(1-\delta\right)\right]}\right\}^{\frac{1}{\gamma}} X_t. \quad (24)$$

As $0 \leq \delta \leq 1$ and $0 < \beta < 1$, it is easily shown that the ratio $\tilde{K}_t/K_t^{f*}$ is larger than one:

$$\frac{\tilde{K}_t}{K_t^{f*}} = \left\{1 + \frac{q}{m}\left[1 - \beta\left(1-\delta\right)\right]\right\}^{\frac{1}{\gamma}} > 1. \quad (25)$$

We define $K_t^* = K^*(X_t; \sigma)$ as the optimal capital level under the irreversibility constraint, where only the user cost effect is active. So, for example, when a firm starts without any initial capital, it immediately invests so that the capital stock is equal to K_t^*. However, if the inherited capital from the previous period $(1 - \delta) K_{t-1}$ exceeds K_t^*, the actual capital stock and this optimal capital stock diverge.

We know from previous studies that K_t^* is equal to or lower than K_t^{f*} because of the irreversibility constraint. Thus, we have the following relationship:

$$K_t^* \leq K_t^{f*} < \tilde{K}_t. \tag{26}$$

Since

$$\frac{\tilde{K}_t}{K_t^*} \geq \frac{\tilde{K}_t}{K_t^{f*}} \ \ (> 1), \tag{27}$$

$\tilde{K}_t/K_t^{f*}$ serves as the floor for $\tilde{K}_t/K_t^*$.

The degree of variation in the capacity utilization rate U_t is determined by the ratio of $\tilde{K}_t$ to K_t^*. When the $\tilde{K}_t$ is too high in proportion to K_t^*, the capacity utilization rate always hits the ceiling of unity. In order for capacity utilization to vary along the realization of X_t, the ratio of $\tilde{K}_t$ to K_t^* should be close to one. As Eq.(25), the ratio of $\tilde{K}_t/K_t^{f*}$, which constitutes the floor for $\tilde{K}_t/K^*$, is negatively related to the ratio of m/q.

The relationship between the degree of variation in capacity utilization U_t and the degree of uncertainty σ is unclear. A larger σ makes K_t^* smaller due to the user cost effect, which leads to a larger $\tilde{K}_t/K_t^*$. At the same time, a larger σ raises the probability of a sudden shock in the stochastic demand shift factor X_t. A large negative shock leads to a smaller $\tilde{K}_t$, which activates the hangover effect. The total effect is therefore ambiguous.

3.3 Model with varying capacity utilization (Leontief production function)

We provide an alternative model without any maintenance costs. To let the capacity utilization fluctuate, we instead assume a Leontief production function where labor and capital are combined at a fixed proportion.

The demand function is isoelastic as in the previous models:

$$Q_t = X_t P_t^{-\epsilon}. \tag{28}$$

However, the production technology is now of a Leontief type:

$$Q_t = \min\left(L_t, \psi K_t^a\right), \tag{29}$$

where ψ denotes a fixed parameter. Note that it is active capital, not the capital stock, that affects production. The capital stock evolves according to Eq.(8).

Because labor and capital are combined at a fixed proportion, the active capital K_t^a depends on labor input. At the same time, they are both constrained by the current capital stock K_t:

$$K_t^a = \frac{1}{\psi}L_t \leq K_t. \tag{30}$$

Since labor is flexible, the optimal level of L_t is determined by maximizing the operating profit $\pi_t = P_t Q_t - w L_t$. However, as Eq.(30) shows, L_t is limited by ψK_t. Therefore,

$$L^* = \min\left(\left(\frac{\epsilon - 1}{\epsilon w}\right)^\epsilon X_t, \ \psi K_t\right). \tag{31}$$

The firm's dynamic optimization problem is rewritten as

$$V\left(K_{t-1}, X_t\right) = \max_{I_t \geq 0} X_t^{\frac{1}{\epsilon}} \left(L_t^*\right)^{1-\frac{1}{\epsilon}} - w L_t^* - q I_t + \beta E_t\left[V\left(K_t, X_{t+1}\right)\right]. \tag{32}$$

Again, the solution is expressed by

$$I_t = I\left(K_{t-1}, X_t; \sigma\right). \tag{33}$$

In a special case where there is no irreversibility constraint, the optimal capital level in each period is

$$K_t^{f*} = \frac{1}{\psi}\left(\frac{\epsilon - 1}{\epsilon w}\right)^\epsilon X_t. \tag{34}$$

4   Numerical solution to investment functions

We now solve for optimal investment

$$I_t = I\left(K_{t-1}, X_t; \sigma\right) \tag{35}$$

for the different model settings discussed above, using numerical dynamic programming methods. The obtained functions will be used to compare the short-run and the long-run effects of uncertainty on the firm's investment, capital use, and labor input in these models.

In all cases, the deterministic trend μ is fixed at 0.029, the discount factor β is set to 0.9523 ($= 1/(1 + r)$ where $r = 0.05$), the price elasticity of demand ϵ is set to 10 and there is no capital depreciation, i.e., $\delta = 0$, as in Abel and Eberly (1999). For the maintenance cost model, we set the cost share of capital in the Cobb-Douglas production function α to 0.33. All the other parameters can be arbitrarily chosen (except for the ratio between m and q) since they do not affect our results.

Regarding the maintenance cost model, the variation in capacity utilization depends on the ratio of m to q. We assume two different ratios, $m/q = 1$ and $m/q = 0.25$, to see how the ratio affects the result. As we see in Eq.(25), the lower ratio is expected to generate a larger variation in capacity utilization.

Regarding the Leontief model, we set ψ to the value at which the optimal cost share of capital is equal to 0.33 as in the Cobb-Douglas models.

We assume that the firm is risk neutral. We initially set the value function $V(K_{t-1}, X_t)$ to some value (typically zero) and, at each iteration, we choose the optimal investment value to maximize the value function, depending on the state variables K_{t-1} and X_t and the expected value of the value function in the next period $V(K_t, X_{t+1})$.

4.1 Short-run investment dynamics

To highlight the firm's short-run investment behavior over the business cycle, we generate a wave-like sample path of the demand shock X_t where the standard deviation σ is set to 0.24. The development paths of the firm's capital stock in the different models are presented in Figure 1. Since the parameters, forms of the functions, and the associated capital user costs differ across the models, we normalize the paths with the capital stock path in the respective frictionless model.

[Figure 1: Sample paths of the capital stock (K) and active capital (K^a)]

In the frictionless model, the firm immediately adjusts its capital stock both upwards and downwards to the optimal level where the marginal revenue product of capital equals the Jorgensonian user cost.

In the AE model with full irreversibility, the firm makes a new investment only when the marginal revenue product of capital exceeds a significantly higher hurdle because of the user cost effect due to the irreversibility constraint. As a result, the firm's capital stock responds less to a positive demand shock than in the frictionless model.

In the two maintenance cost models, the firm has a lower investment hurdle and thus invests more in economic booms than what is the case in the AE model. The user cost of capital in the maintenance cost model consists of both the purchase cost q and the maintenance cost m as described in Eq.(22).

The firm is able to put aside some redundant capital stock in a recession and thereby avoid paying the maintenance cost for that part of its capital. Hence, the effective user cost of capital is lower the more the demand shock fluctuates (i.e., a larger σ). This explains why the firm in the maintenance cost models tend to invest more in economic booms. This also answers the question of why the firm invests more with $m/q = 1$ than with $m/q = 0.25$, as the share of the maintenance cost in the capital user cost is larger in the former case.

In a recession, on the other hand, capacity utilization is lower than unity in the maintenance cost models because the firm has an incentive to cut the operating cost by reducing capital use. At the same time, more capital is active in the maintenance cost models than in the frictionless model. This is because the firm in the maintenance cost models only needs to equate the marginal revenue product of capital with m when it has excessive capacity, while the firm in the frictionless model needs to equate the marginal revenue product of capital with the whole user cost of capital. Note also that, because the maintenance cost is relatively lower in the case with $m/q = 0.25$ than in the case with $m/q = 1$, the firm has higher capacity utilization in the former case.

The Leontief model yields a similar result as the maintenance cost models. The amount of active capital is determined by labor input as long as it does not exceed the capital stock. Capacity utilization decreases when the firm faces a negative demand shock.

4.2   Short-run effect on labor

In Figure 2, we illustrate how labor input varies in the different models over the sample business cycle generated in the previous subsection. Figure 3 shows the associated ratio between labor and active capital.

[Figure 2: Sample paths of labor (L)]

[Figure 3: Factor ratios between labor (L) and active capital (K^a)]

In the frictionless model, labor demand goes hand in hand with the capital stock because the ratio of wage to the capital user cost is constant.

In the AE model with full irreversibility, labor demand L_t is determined by the capital stock K_t and the demand shock X_t as in Eq.(4). Since the firm is unable to cut the capital stock in a recession, labor demand is reduced only due to a reduction in X_t. And as the capital stock stays at the same level even in a boom in this example, labor demand increases only due to the rise in X_t. The fluctuation of labor demand is therefore small.

In the maintenance cost models, labor demand L_t is determined by active capital K_t^a and the demand shock X_t as in Eq.(15). Labor demand therefore decreases with active capital in a recession. In a boom, on the other hand, labor demand increases even when active capital K_t^a has reached the capacity limit K_t, simply because X_t is still rising.

There are two regimes with respect to the relationship between the marginal labor cost and the marginal capital cost. In one regime, where the capacity utilization rate is below one (typically in a recession), the ratio of the marginal cost is w/m because the maintenance cost m is the only determinant of active capital. In the other regime, where the capacity utilization rate has reached unity (typically in a boom), the ratio of the marginal costs is w/c because the firm needs to expand the capital stock with a new investment. Since $c > m$, the relative labor cost is lower in the second regime. This can be interpreted as a substitution in factor demand from capital to labor. The shift of the cost ratio and the labor-capital substitution is easily understood in Figure 3, where the factor demand ratio of labor to capital increases in a boom while it stays at a fixed ratio in a recession.

In the Leontief model, L_t and K_t^a are jointly determined by X_t. The ratio between labor and active capital is constant. In a recession, labor demand decreases together with active capital. In a boom, however, once capital use reaches full capacity, labor demand no longer increases unless new capital investments are made to expand the capacity.

4.3 Long-run effect on the capital stock

We examine in this subsection the effect of the different model assumptions on capital accumulation and capital use in the long run. For this purpose, we analyze the ratio between the expected level of the capital stock under the irreversibility constraint and that in the respective frictionless case for each model (i.e., the AE model, the two maintenance cost models, and the Leontief model).

We define $\kappa_K(t)$ as the ratio of the expected level of the capital stock at date t under full irreversibility $E[K_t]$ to that in the frictionless case $E\left[K_t^f\right]$:

$$\kappa_K(t) \equiv \frac{E[K_t]}{E\left[K_t^f\right]}. \tag{36}$$

Note that $E\left[K_t^f\right]$ is constant whereas $E[K_t]$ varies with σ.

Similarly, we define $\kappa_{K^a}(t)$ as the ratio of the expected level of active capital at date t under full irreversibility to that in the frictionless case. However, as the capital stock and the active capital stock are equal in the frictionless cases, $\kappa_{K^a}(t)$ is expressed as

$$\kappa_{K^a}(t) \equiv \frac{E[K_t^a]}{E\left[K_t^f\right]}. \tag{37}$$

The long-run capacity utilization rate is defined as

$$\kappa_U(t) \equiv \frac{E[K_t^a]}{E[K_t]} = \frac{\kappa_{K^a}(t)}{\kappa_K(t)}. \tag{38}$$

To obtain the expected levels of the capital stock and active capital, we simulate capital accumulation and capital use for n hypothetical firms, each facing a random realization of the demand shock X_t according to geometric Brownian motion Eq.(2). n is set to 10 million for a small value of σ (< 0.16), while it is set to 40 million for a larger value of σ (≥ 0.16). t is set to 100 because we believe that 100 years is long enough to eliminate any effect of the initial capital stock, so that we can approximate the expected levels of capital accumulation and capital use in an infinite time horizon ($t = \infty$). We take the averages of the capital stock and active capital at $t = 100$ to obtain $\kappa_K(t)$ and $\kappa_{K^a}(t)$.

Figures 4–7 illustrate the values of $\kappa_K(t)$ and $\kappa_{K^a}(t)$ under the four different model settings, over the range from $\sigma = 0$ to $\sigma = 0.24$.

[Figure 4: $\kappa_K(t)$ for the AE model]

[Figure 5: $\kappa_K(t)$ and $\kappa_{K^a}(t)$ for the maintenance cost model $(m/q = 1)$]

[Figure 6: $\kappa_K(t)$ and $\kappa_{K^a}(t)$ for the maintenance cost model $(m/q = 0.25)$]

[Figure 7: $\kappa_K(t)$ and $\kappa_{K^a}(t)$ for the Leontief model]

$\kappa_K(t)$ in Figure 4 is a numerical replication of $\kappa(\infty)$ in Figure 1 in Abel and Eberly (1999). It fits Abel and Eberly's $\kappa(\infty)$ fairly well. This fact confirms the appropriateness of our numerical method.

Figure 5 plots $\kappa_K(t)$ and $\kappa_{K^a}(t)$ against σ in the maintenance cost model with $m/q = 1$. The capital stock $\kappa_K(t)$ increases with uncertainty, while active capital $\kappa_{K^a}(t)$ decreases. This implies that the hangover effect dominates in the whole studied range of σ and that capacity utilization $\kappa_U(t)$ falls with σ. We also observe that the overinvestment becomes more prominent at a larger σ. These findings can be explained by two factors. In the maintenance cost model, the user cost of capital consists of both the purchase cost q and the maintenance cost m as in Eq.(22). The firm is able to put aside some redundant capital stock in a recession and thereby avoid paying the maintenance cost for that part of its capital. Hence, the effective user cost of capital decreases when the probability that the firm faces a recession increases (i.e., at higher levels of uncertainty).

Another factor is labor costs. The reluctance toward new investments is not only due to the capital costs that are sunk, but also to the labor costs that depend on active capital. Because the firm is able to reduce active capital and thereby labor input, the total binding costs of new investments, including the user cost of capital and the labor cost, is less than in the AE model. Moreover, its expected value becomes smaller when the firm is more likely to face a recession.

Active capital $\kappa_{K^a}(t)$ is decreasing with σ, and the reduction is larger than in the case with the AE model (remember that $\kappa_K(t) = \kappa_{K^a}(t)$ in the AE model). This finding is explained in the same way: because capacity utilization can now be lower than unity and use of capital is costly, the firm makes less use of the capital stock than what was the case in the AE model.

Figure 6 presents $\kappa_K(t)$ and $\kappa_{K^a}(t)$ in the maintenance cost model with $m/q = 0.25$. As in the case with $m/q = 1$, the capital stock $\kappa_K(t)$ increases, the active capital stock $\kappa_{K^a}(t)$ decreases, and capacity utilization $\kappa_U(t)$ also decreases with uncertainty σ. In this model, however, the maintenance cost m makes up a smaller part of the whole user cost of capital. This means that the

flexible part of the user cost is smaller. The actual user cost does not decrease as much as in the model with $m/q = 1$. This explains why the firm accumulates less capital.

As discussed above (see Eq.(25)), the variation in the capital utilization rate is expected to increase with the ratio m/q. This is confirmed by the result that the capacity utilization rate at $\sigma = 0.24$ in the model with $m/q = 1$ is 73% while the capacity utilization rate at $\sigma = 0.24$ in the model with $m/q = 0.25$ is 86%.

Figure 7 illustrates $\kappa_K(t)$ and $\kappa_{K^a}(t)$ in the the Leontief model. The capital stock $\kappa_K(t)$, the active capital stock $\kappa_{K^a}(t)$, and capacity utilization $\kappa_U(t)$ react in the same direction as in the maintenance cost models discussed above. The degree of the reaction is in between the two models. It is interesting to find a similar result even in the absence of maintenance costs.

Figure 8 highlights comparison of $\kappa_{K^a}(t)$ for the four models. Figure 9 plots the associated capacity utilization rates against σ.

[Figure 8: Comparison of $\kappa_{K^a}(t)$ for the different models]

[Figure 9: Comparison of U_t for the different models]

4.4 Long-run effect on labor demand

Exactly as the analysis of capital stock, we define $\kappa_L(t)$ as the ratio between the expected level of labor demand at date t with the capital irreversibility constraint,[6] $E[L_t]$, and that in the frictionless case, $E\left[L_t^f\right]$:

$$\kappa_L(t) = \frac{E[L_t]}{E\left[L_t^f\right]}. \tag{39}$$

Figure 10 depicts $\kappa_L(t)$ in the different models. In all cases, $\kappa_L(t)$ moves in a similar way and follows $\kappa_K(t)$. Note the absence of an initial surge in $\kappa_L(t)$ in the AE model. Remember that $\kappa_K(t)$ initially rise before a sharp decline. However, since labor is flexible unlike the capital stock, the hangover effect of capital does not affect labor demand very much.

[Figure 10: Comparison of $\kappa_L(t)$ for the different models]

---

[6]Note that labor itself is always flexible.

5 Conclusion and discussion

This study has analyzed the effects of uncertainty on capital and labor use in both the short run and the long run in models where firms have incentives not to operate at full capacity. We modify the real options model (the AE model) proposed by Abel and Eberly (1999) to introduce a capital "maintenance cost." The firm needs to pay this cost for all capital used for production, in addition to the capital purchase cost it pays when acquiring the capital stock. This cost makes it costly for firms to operate at full capacity in every period. We also proposes another model setting for variable capacity utilization. In this model, the production function is a Leontief instead of a Cobb-Douglas. Labor and capital are combined at a fixed ratio. Since labor, which is assumed to be flexible, is costly to use, both labor and capital are used less when firms face negative demand shocks. For each of these model settings, we estimate the dynamic investment function for given values of the state variables (the capital stock and demand shock) and uncertainty through numerical simulation. Based on these investment functions, we analyze how firms behave in terms of capital accumulation as well as capital and labor use, in both the short run and the long run. In the analysis, we use the predictions by the AE model as a benchmark.

In the short run, the firm in our maintenance cost models accumulates more capital in economic booms compared to the firm in the AE model. This is because the effective user cost of capital, which consists of the maintenance cost and the purchase cost, decreases. It also decreases (hence more capital is accumulated) when uncertainty increases.

However, because the user cost still exceeds the traditional Jorgensonian capital user cost, the firm does not invest as much as in the frictionless case. Reluctance toward new investments rises when the capital maintenance cost becomes smaller relative to the capital purchase cost.

In a recession, on the other hand, the firm in our maintenance cost models is able to reduce active capital. Compared to the frictionless case, the amount of active capital is larger. This is because it is cheap to raise the production level by utilizing more capital when the firm has already built an excessive capital stock, compared to the frictionless case where the firm needs to pay a purchase cost to expand the capital stock. We also find that the activity level of capital is higher when the maintenance cost is lower relative to the purchase cost.

Regarding labor input, our study shows that the firm in the AE model will not reduce as much labor input in a recession as the firm in our maintenance

cost models because capital input is fixed. In contrast, labor input in the maintenance cost models falls with capital input. It is worth noting that there is a demand shift in favor of labor when the capacity limit of capital has been reached. The firm prefers labor input to new investment because the former is relatively cheaper.

The long-run effect of uncertainty on capital accumulation in our maintenance cost model is that the firm tends to build a larger capital stock as uncertainty increases. This is explained by the fact that the effective user cost of capital as well as the binding cost of labor related to new investments become lower. Another interesting result is that the firm's active capital decreases as uncertainty increases, although capital accumulation at the same time intensifies. This implies that the expected capacity utilization is a decreasing function of uncertainty.

Labor demand decreases together with active capital when uncertainty increases, except at low levels of uncertainty. An interesting feature is that labor demand in the AE model also decreases though the capital accumulation initially rises.

The Leontief model yields similar results to the ones from the maintenance cost models.

We expect our proposed models to better predict capital accumulation, capital input, and labor demand. Empirical studies are necessary to test their validity.

References

[1] Abel, Andrew B. (1981), "A Dynamic Model of Investment and Capacity Utilization," *The Quarterly Journal of Economics*, 96(3), 379-403.

[2] Abel, Andrew B. (1983), "Optimal Investment under Uncertainty," *American Economic Review*, 73(1), 228-33.

[3] Abel, Andrew B. (2002), "On the Invariance of the Rate of Return to Convex Adjustment Costs," *Review of Economic Dynamics*, 5(3), 586-601.

[4] Abel, Andrew B. and Eberly, Janice C. (1996), "Optimal Investment with Costly Reversibility," *Review of Economic Studies*, 63(4), 581-93.

[5] Abel, Andrew B. and Eberly, Janice C. (1999), "The effects of irreversibility and uncertainty on capital accumulation," *Journal of Monetary Economics*, 44(3), 339-77.

[6] Bernstein, Jeffrey I. and Nadiri, M. Ishaq (1988), "Investment, Depreciation, and Capital Utilization," NBER Working Papers 2571, National Bureau of Economic Research.

[7] Bertola, Guiseppe and Caballero, Ricardo J. (1994), "Irreversibility and Aggregate Investment," *Review of Economic Studies*, 61(2), 223-46.

[8] Bischoff, Charles W. and Kokkelenberg, Edward C. (1987), "Capacity Utilization and Depreciation-in-Use," *Applied Economics*, 19(8), 995-1007.

[9] Bloom, Nicholas (2009), "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623-85.

[10] Bloom, Nicholas; Bond, Stenphen R.; and van Reenen, John (2003), "Uncertainty and Company Investment Dynamics: Empirical Evidence for UK Firms," CEPR Discussion Papers 4025, Centre for Economic Policy Research.

[11] Bloom, Nicholas; Bond, Stenphen R.; and van Reenen, John (2007), "Uncertainty and Investment Dynamics," *Review of Economic Studies*, 74(2), 391-415.

[12] Bond, Stenphen R. and Lombardi, Domenico (2006), "To Buy or Not to Buy? Uncertainty, Irreversibility, and Heterogeneous Investment Dynamics in Italian Company Data," *IMF Staff Papers*, 53(3), 2.

[13] Bond, Stephen R.; Söderbom, Måns; and Wu, Guiying (2011), "Pursuing the wrong options? Adjustment costs and the relationship between uncertainty and capital accumulation," *Economics Letters*, 111(3), 249-51.

[14] Chirinko, Robert S. and Schaller, Huntley (2009), "The irreversibility premium," *Journal of Monetary Economics*, 56(3), 390-408.

[15] Christiano, Lawrence J.; Eichenbaum, Martin; and Evans, Charles L. (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1-45.

[16] Cooper, Russell W. and Haltiwanger, John (2006), "On the nature of capital adjustment costs," *Review of Economic Studies*, 73(3), 611-34.

[17] Dixit, Avinash and Pindyck, Robert S. (1994), *Investment under uncertainty*, Princeton University Press.

[18] Fousekis, Panos and Stefanou, Spiro E. (1996), "Capacity Utilization under Dynamic Profit Maximization," *Empirical Economics*, 21(3), 335-59.

[19] Guiso, Luigi and Parigi, Giuseppe (1999), "Investment and Demand Uncertainty," *The Quarterly Journal of Economics*, 114(1), 185-227.

[20] Leahy, John V. and Whited, Toni M. (1996), "The Effect of Uncertainty on Investment: Some Stylized Facts," *Journal of Money, Credit and Banking*, 28(1), 64-83.

[21] Morrison, Catherine J. (1986), "Productivity measurement with non-static expectations and varying capacity utilization: An integrated approach," *Journal of Econometrics*, 33(1-2), 51-74.

[22] Pindyck, Robert S. (1988), "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review*, 78(5), 969-85.

[23] Pindyck, Robert S. (1993), "A Note on Competitive Investment under Uncertainty," *American Economic Review*, 83(1), 273-77.

[24] Prucha, Ingmar R. and Nadiri, M. Ishaq, (1996), "Endogenous capital utilization and productivity measurement in dynamic factor demand models Theory and an application to the U.S. electrical machinery industry," *Journal of Econometrics*, 71(1-2), 343-79.

[25] Smith, Kenneth R. (1969), "The effect of uncertainty on monopoly price, capital stock and utilization of capital," *Journal of Economic Theory*, 1(1), 48-59.

[26] Smith, Kenneth R. (1970), "Risk and the Optimal Utilization of Capital," *Review of Economic Studies*, 37(2), 253-59.

[27] Taubman, Paul and Wilkinson, Maurice (1970), "User Cost, Capital Utilization and Investment Theory," *International Economic Review*, 11(2), 209-15.
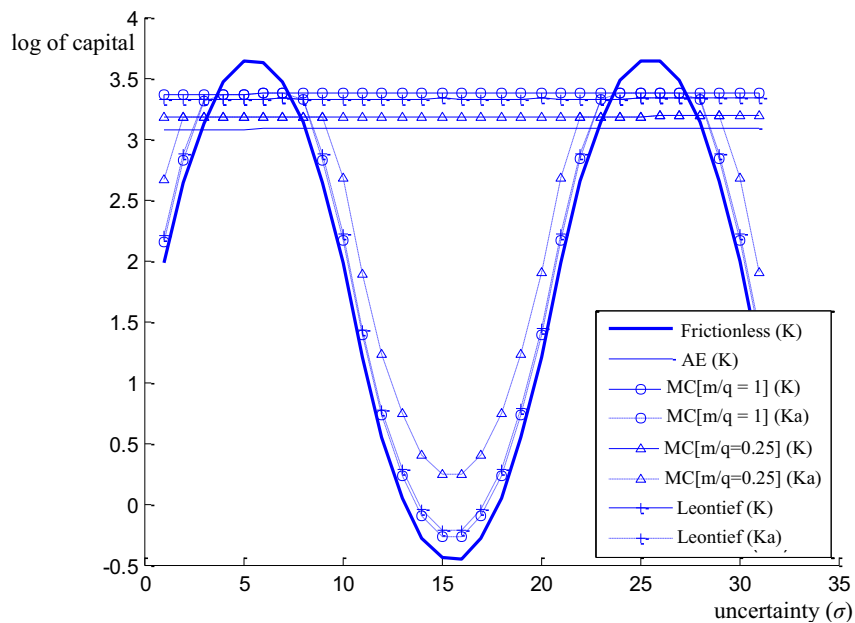
Figure 1: Sample paths of the capital stock (K) and active capital (Ka)
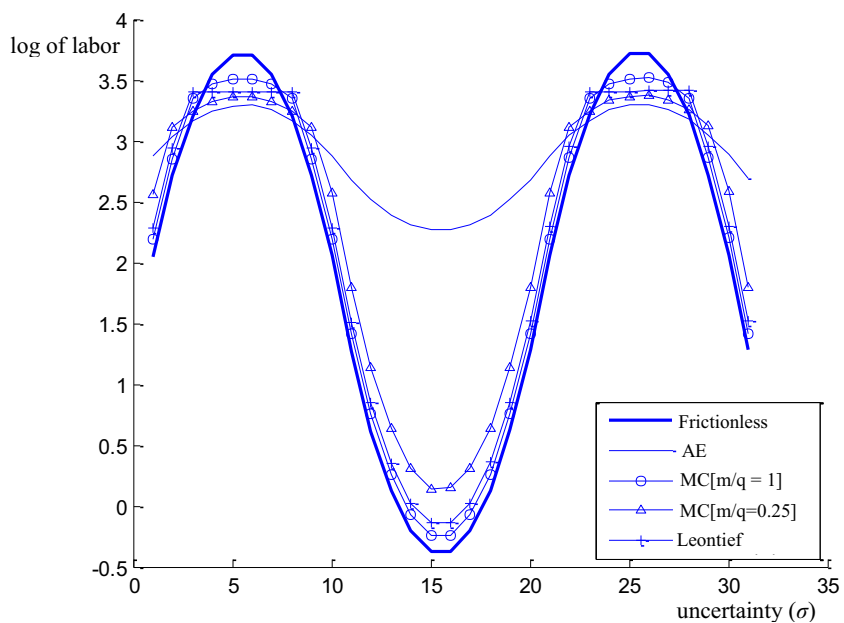(Note: AE stands for the Abel-Eberly model, MC stands for the maintenance cost model)
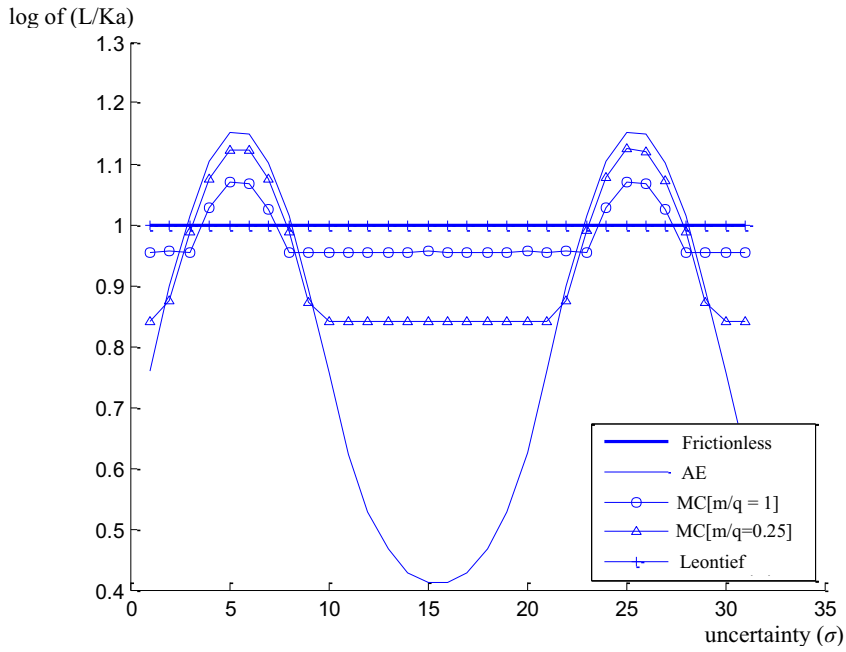

Figure 2: Sample paths of labor (L)

Figure 3: Factor input ratio L/Ka (i.e., the ratio of labor to active capital)

Figure 4: $\kappa_K(t)$ for the AE model
(Note: $\kappa_K(t)$ is the ratio of the expected level of the capital under full irreversibility to that in the frictionless case.)
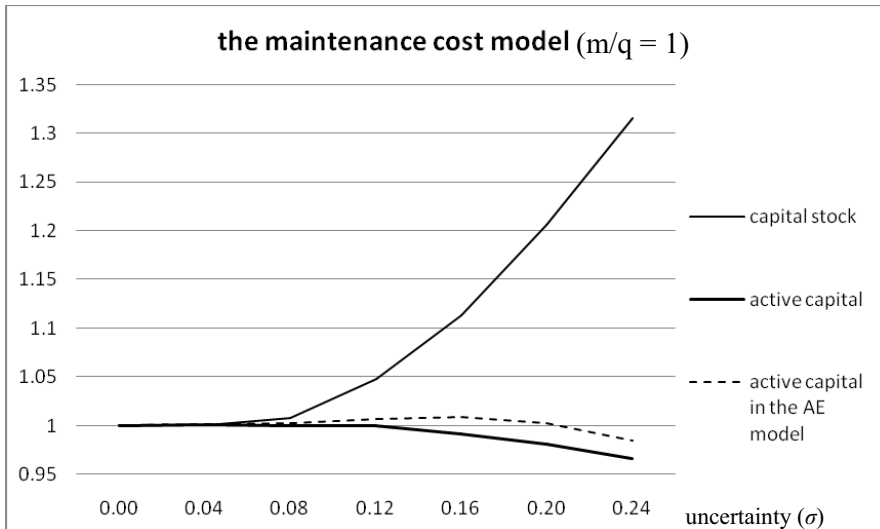
Figure 5: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ for the maintenance cost model (m/p = 1)
(Note: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ are the ratios of the expected levels of the capital stock and active capital under full irreversibility to those in the frictionless case, respectively.)
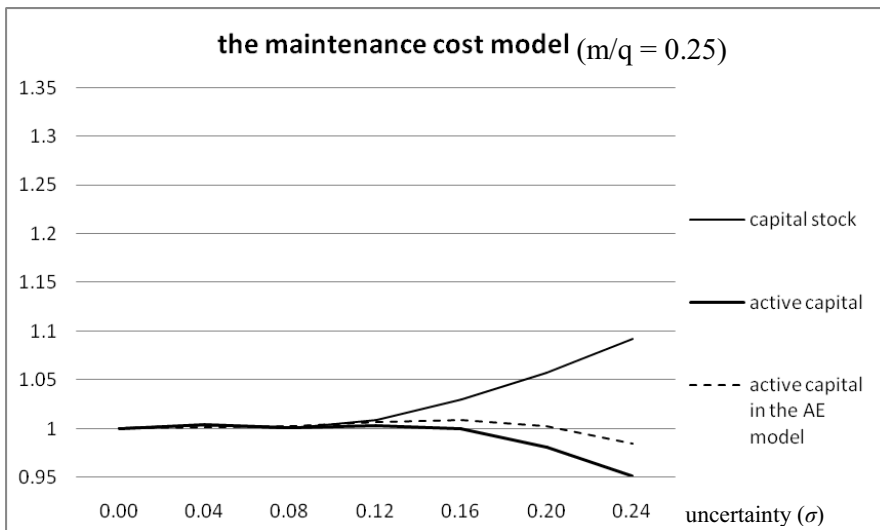
Figure 6: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ for the maintenance cost model (m/p = 0.25)
(Note: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ are the ratios of the expected levels of the capital stock and active capital under full irreversibility to those in the frictionless case, respectively.)
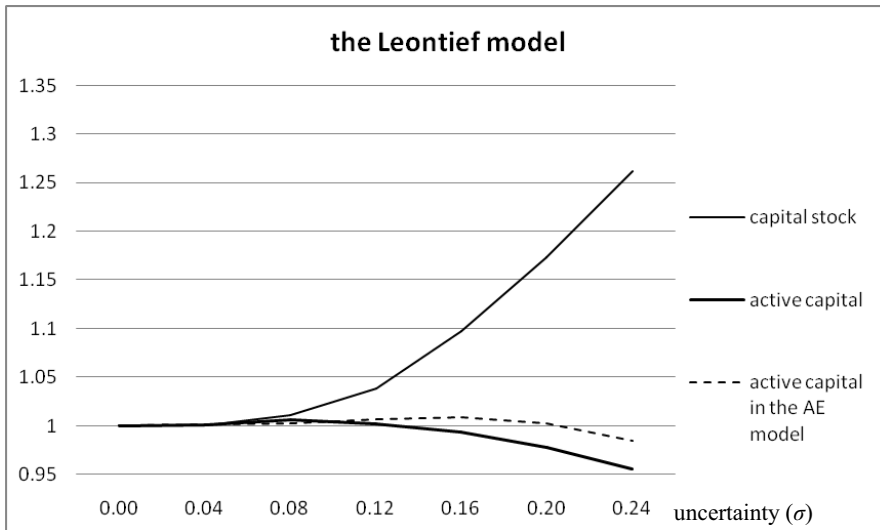
Figure 7: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ for the Leontief model
(Note: $\kappa_K(t)$ and $\kappa_{Ka}(t)$ are the ratios of the expected levels of the capital stock and active capital under full irreversibility to those in the frictionless case, respectively.)
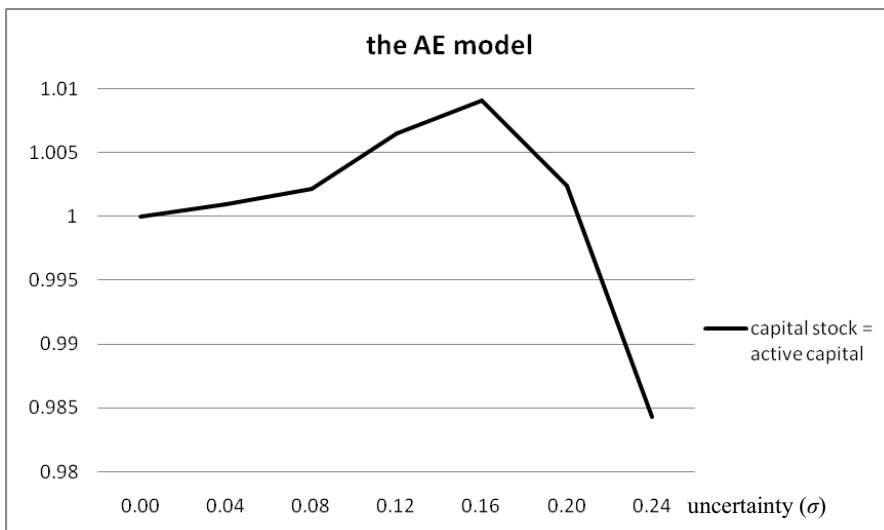

Figure 8: Comparison of $\kappa_{Ka}(t)$ for the different models
(Note: $\kappa_{Ka}(t)$ is the ratio of the expected level of active capital under full irreversibility to that in the frictionless case.)

Figure 9: Comparison of U_t between the different models
(Note: U_t is the expected long-run capacity utilization rate.)

Figure 10: Comparison of $\kappa_L(t)$ for the different models
(Note: $\kappa_L(t)$ is the ratio of the expected level of labor demand under full irreversibility to that in the frictionless case.)
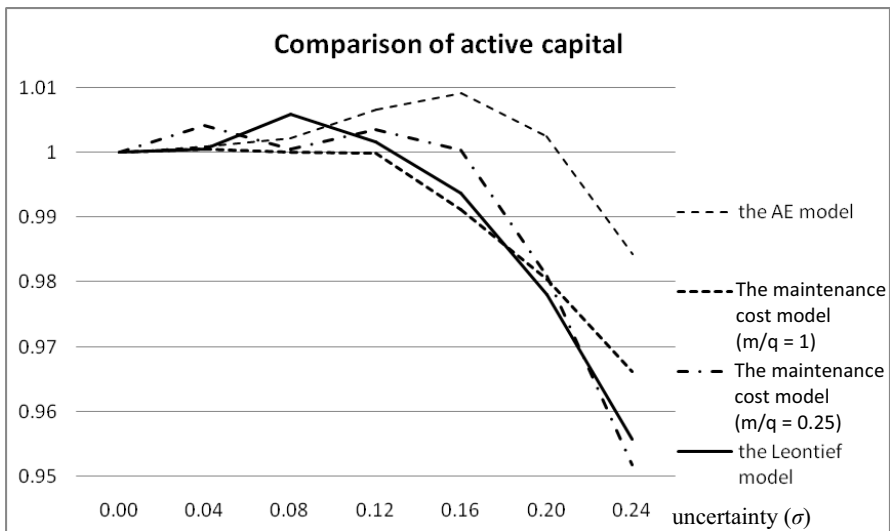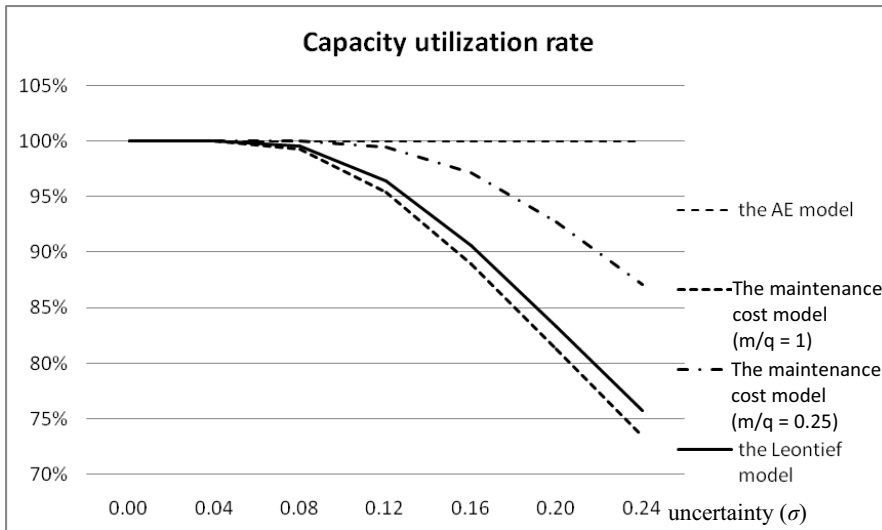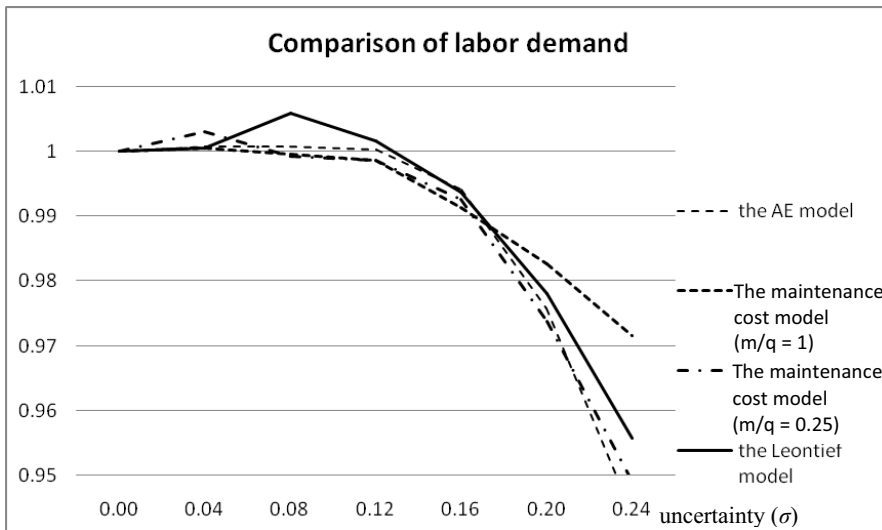
Paper 2

Capital Adjustment under Variable Capacity Utilization

Måns Söderbom and Yoshihiro Sato
University of Gothenburg

Abstract

This study generalizes the existing real options model so as to accommodate fluctuations in capacity utilization, and then analyzes the effects of uncertainty in this modified model. Using both simulated data and Swedish 2-digit sectoral data for the manufacturing industry, we show that the main argument of the real options model – that the responses of the capital stock to demand shocks are weaker at higher levels of uncertainty – still holds, implying that policy stimuli have limited effects on investment in the immediate aftermath of an uncertainty shock. On the other hand, we find that actual capital use (active capital) flexibly responds to demand shocks through adjustments of capacity utilization even at high levels of uncertainty, which suggests that policy stimuli have positive impacts on production activities even at high uncertainty.

1 Introduction

The real options literature in the capital investment theory analyzes the effects of uncertainty on investment dynamics when firms are constrained by total or partial irreversibility of investments. Theoretical studies such as Abel and Eberly (1999) and Caballero (1999) show that responsiveness of investments to demand shocks are weaker when firms face higher levels of uncertainty. This so-called "real options effect" arises because firms regard it advantageous to delay an investment decision. The reason is that capital is hard to adjust downwards due to irreversibility constraints when arrival of new information changes the firm's optimal capital stock level. Bloom, Bond, and van Reenen (2003) empirically test the real options effect using both simulated data and firm-level data from the U.K. and find evidence in support of the real options model.

The real options effect has important policy implications. Bloom et al.(2003) argue, based on their empirical findings, that responses of investments to a policy

stimulus such as monetary or fiscal policy may be seriously limited at higher levels of uncertainty. Similarly, Bloom (2009) argues that a larger real options value of inaction (i.e., zero gross investment) at higher levels of uncertainty makes firms more cautious, which leads to limited short-run effects of a policy stimulus in the immediate aftermath of an uncertainty shock.

In the existing model mentioned above, it is assumed that firms always operate at full capacity, i.e., that they make full use of the capital stock once an investment is made. This is because firms have no incentive to use less capital than the installed capital stock in the model. Consequently, capital is treated as quasi-fixed, and capital adjustment is defined as changes in the capital stock by new gross investments and depreciation. In practice, however, we observe pro-cyclical variation in capacity utilization. The quarterly data on capacity utilization, available from Statistics Sweden and the National Institute of Economic Research (NIER) in Sweden, reveals that the capacity utilization rate fluctuates considerably over the business cycle, as shown later in our study. This fact implies that actual capital use deviates from the installed capital stock, and suggests that the capital stock may rather be flexible internally in the sense that the firm is able to adjust its capital input by changing its capacity utilization.

The present study generalizes the real options model in Abel and Eberly (1999), hereafter called "the AE model", so as to accommodate fluctuations in capacity utilization, and analyzes the effects of uncertainty on investment and production activities in this modified model. We also test empirically its validity. The model we propose assumes a profit-maximizing firm that is able and may have an incentive to adjust its capacity utilization of the capital stock. For more details of the model, see Chapter 1. In the model, capital is adjusted in two ways: adjustment of the capital stock (i.e., investment) as in the AE model, and adjustment of actual use of the capital stock, which we call "active capital".

For empirical investigation, we use real data for Swedish manufacturing industries at the 2-digit sectoral level during the period 1990–2010. Data on capacity utilization and our empirical measures on uncertainty are only available at the 2-digit sectoral level from Statistic Sweden and NIER. Regarding uncertainty, we define it as the gap between firms' expectations for demand shocks and their realization. Imagine that the output demand for a firm increase by 10 percent. If the firm has expected a demand increases by exactly 10 percent, there is no surprise for the firm. In contrast, if the firm has expected a

demand increase by 20 percent or no demand increase, this is a misjudgment by the firm. We assume that this gap reflects uncertainty in the market. Note that this measure of uncertainty is separated from demand shocks. The reaction of the firm may be different depending on uncertainty even if the change in output demand is the same. To construct the empirical measure on uncertainty, we use the Economic Tendency Survey by NIER where the net figures on firms' expectations of demand shifts and the realization are available at the 2-digit sectoral level.

We define linear empirical equations to test the validity of the structural model proposed in Chapter 1. However, it is not possible to tell whether the signs of the estimated coefficients in the empirical equations based on data generated from the structural model are as we expect. Neither is it possible to tell whether the coefficients estimated based on the firm-level data are consistent with those estimated based on the sector-aggregated data due to aggregation biases. In other words, it is not clear whether the firm-level mechanism for decision-making regarding investment and capital use can be detected in econometric analyses using aggregated data at the sectoral level. Therefore, we first let the structural model with a suitable set of parameters generate series of data at the firm level and estimate the empirical equations. Then, we aggregate the simulated data to the sectoral level and estimate the empirical equations to investigate whether the estimated coefficients have the same signs as those estimated using the firm-level data. Finally, we estimate the empirical model using real data at the sectoral level.

The empirical investigation finds that the adjustment of the capital stock is slow and that it is approximated by a convex function of demand shifts. It also shows that the effect of uncertainty on the adjustment is, if anything, negative. This is consistent with the prediction by the real options AE model and the empirical findings in Bloom et al.(2003), implying that the effects of uncertainty are not quantitatively different when allowing for variable capacity utilization. However, we find that the effect of uncertainty is, if anything, positive on the adjustment of both active capital and capacity utilization. The results suggest that firms tend to vary capacity utilization more at higher levels of uncertainty. This can be explained in two ways: (i) firms prefer internal adjustment by varying their capacity utilization over adjusting the capital stock at a high level of uncertainty, and (ii) when uncertainty is large, firms tend to sit with vacant capacity (a so-called "hangover" effect; see Abel and Eberly, 1999) and easily adjust their active capital in response to positive demand shocks.

Our study sheds new light on short-run investment dynamics and "rigidity" in the capital adjustment process. We empirically show that firms in fact flexibly adjust their active capital to demand shocks. This means that production is also adjusted flexibly to demand shocks. This finding suggests that, while policy stimuli such as monetary or fiscal policy still have a limited impact on investment activities at higher levels of uncertainty, they have a considerable impact on production activities.

Section 2 discusses the main features of the real options model, and generalizes it so as to accommodate fluctuations in capacity utilization. Section 3 describes the Swedish data used in the empirical analysis and highlights significant variation in the capacity utilization rate observed in the data. Section 4 presents the empirical specification based on an error correction model. Sections 5 and 6 discuss the estimation results for simulated data and the real data, respectively, followed by a conclusion in Section 7.

2   Conceptual framework

2.1   The real options theory

The capital investment literature highlights the importance of partial or total irreversibility and uncertainty; see, e.g., Caballero (1991, 1999), Ingersoll and Ross (1992), Abel and Eberly (1994, 1996, 1999), Bertola and Caballero (1994), Dixit and Pindyck (1994), Bloom (2000, 2007, 2009), Bloom et al.(2003, 2007), Bigsten et al.(2005), and Bond et al.(2011).[1] A combined effect of uncertainty and irreversibility generates an "options value" in the sense that a firm regards it advantageous to delay its investment decision. Arrival of new information may change the firm's optimal capital stock level. While increasing the capital stock through additional investments may be relatively easy, reducing the capital stock by disinvestment may often be many times difficult as rental and secondary markets are far from perfect. Consequently, the investment behavior becomes more cautious and the firm responds less to demand shocks. The firm has periods of inaction (i.e. zero gross investment) where demand shocks are too small to motivate a change in the capital stock. Cautiousness increases at higher

---

[1] These studies are based on earlier theoretical contributions on irreversibility by Lucas and Prescott (1971), Hartman (1972), Pindyck (1982, 1988), Abel (1983), Bernanke (1983), Brennan and Schwartz (1985), McDonald and Siegel (1986), Majd and Pindyck (1987), and Dixit (1989). Although Arrow (1968) and Nickell (1974) did not explicitly mention irreversibility, they did address the problem of uncertainty.

levels of uncertainty as the "options value" gets larger.

Figure 1 illustrates the real options effect of uncertainty on investment. The firm will undertake investment only when the marginal revenue product of capital lies above the upper threshold, which can be expressed by the Jorgensonian capital user cost of buying capital goods b times the options value relevant for investment ϕ_I (>1). In Figure 1, the upper threshold is denoted A. In a similar manner, the firm will undertake disinvestment only when the marginal revenue product lies below the lower threshold, which can be expressed by the capital user cost of selling capital goods s ($< b$) divided by the option value relevant for disinvestment ϕ_D (>1). In Figure 1, the lower threshold is denoted B. The area between the two thresholds A and B indicates "the region of inaction" where neither investment nor disinvestment is motivated.

[Figure 1: Real options effect of uncertainty on investment]

To empirically test the real options hypothesis, Bloom et al.(2003) use both simulated data and real data on U.K. firms to estimate the following empirical equation based on an error correction model:

$$
\begin{aligned}
\Delta k_{i,t} &= \beta \Delta y_{i,t} + \eta \left(\Delta y_{i,t} \right)^2 + \theta \left(k_{i,t-1} - y_{i,t-1} \right) \\
&\quad + \phi \sigma_{i,t} \Delta y_{i,t} + a_i + b_t + \varepsilon_{i,t},
\end{aligned} \tag{1}
$$

where $k_{i,t}$ denotes the log of the capital stock, $y_{i,t}$ the log of output, $\sigma_{i,t}$ a measure of uncertainty, a_i firm-specific effects, b_i year-specific effects, and $\varepsilon_{i,t}$ is a serially uncorrelated error term. β, η, θ, and ϕ are coefficients. As the left-hand side is approximated to the investment rate *minus* the depreciation rate, this equation is equivalent to an empirical equation for the investment rate. They find positive and significant values for β and η, indicating that the investment is a positive and convex function of demand shocks.[2] Convexity is a result of the region of inaction in Figure 1 as well as possible effects of supermodularity.[3] They also find a negative value for ϕ, indicating that investment responses to demand shocks are smaller at higher levels of uncertainty. This empirical result is consistent with the option values hypothesis.[4]

---

[2] The squared term does not capture the real options effect of negative demand shocks ($\Delta y_{i,t} < 0$). However, Bloom et al.(2003) argues that positive shocks dominate in their dataset.

[3] This means that investment in one type of capital goods triggers investment in other types of capital goods by raising their marginal products.

[4] They also include capital flows and other control variables in some of their specifications.

2.2 Implications of variable capacity utilization

The real options literature focuses on the rigidity of the capital adjustment process. The argument is based on the assumption that, once capital goods are installed, they are in full use until they wear out. Capital is treated as quasi-fixed, and it is assumed that net investment (i.e., gross investment and depreciation) is a main source of the variation in capital input of a firm.

The theoretical model in Chapter 1 relaxes this assumption and suggests a modified structural model where a profit-maximizing firm is able to adjust its capacity utilization of the capital stock. The reason for the firm not to operate at full capacity is that the model assumes some sort of costs related to capital use, in addition to the purchase cost. The additional cost is called "maintenance costs" of capital in use.[5] Assume a Cobb-Douglas production function with two production factors, capital K_t and labor L_t, and assume for simplicity that capital is totally irreversible while labor is fully flexible and that there is no depreciation.[6] The dynamic optimization problem is formulated as

$$V\left(K_{t-1}, X_t, \sigma\right) = \max_{\substack{I_t \geq 0 \\ K_t^a \leq K_t}} \pi\left(K_t^a, X_t\right) - mK_t^a - qI_t + \beta E_t\left[V\left(K_t, X_{t+1}, \sigma\right)\right], \ (2)$$

where $\pi\left(\cdot\right)$ denotes the function for operating profit (i.e., sales *minus* labor costs), K_t the capital stock that evolves according to $K_t = K_{t-1} + I_t$, K_t^a active capital, I_t investment, m the maintenance cost of active capital, q_t the purchase cost of capital, β the discount rate, and E_t an operator for expectation at t where we assume that the firm is risk neutral. X_t is demand shocks, which follows a geometric Brownian motion with mean $\tilde{\mu}t$ and variance $\sigma^2 t$:

$$\ln X_t \equiv x_t = x_{t-1} + \tilde{\mu} + u_t, \quad\quad\quad (3)$$

where $u_t = iid.N\left(0, \sigma^2\right)$ and $\tilde{\mu} \ \left(= \mu - 0.5\sigma^2\right)$ is the growth rate.[7] In this dynamic optimization problem, the state variables are K_t and X_t, and the choice variable is I_t. Although active capital K_t^a is also a choice variable, it is

---

[5] Chapter 1 also proposes another model, which does not require "maintenance costs", where production technology is characterized by a Leontief function. Under this assumption, use of capital and labor depends proportionally on output.

[6] We use the same assumption as in Abel and Eberly (1999) to make the model comparable. We suppose that there is no loss of generality.

[7] For more detail, see Chapter 1.

determined by the static optimization problem

$$K_t^a = \arg \max_{K_t^a \leq K_t} \pi \left(K_t^a, X_t \right) - m K_t^a. \tag{4}$$

The capacity utilization rate CU_t is obtained by $CU_t = K_t^a / K_t$.

By comparison, the existing real options model, which we call the AE model after Abel and Eberly (1999), is characterized by the following optimization problem:

$$V \left(K_{t-1}, X_t, \sigma \right) = \max_{I_t \geq 0} \pi \left(K_t, X_t \right) - q I_t + \beta E_t \left[V \left(K_t, X_{t+1}, \sigma \right) \right]. \tag{5}$$

Because these optimization problems involve non-linear constraints, it is not straightforward to analytically obtain a solution for $I_t = I \left(K_{t-1}, X_t, \sigma \right)$. We use a numerical approach where we give a suitable set of parameters to the model and obtain a solution for I_t through iterations.

The predictions from the theoretical analysis on short-run effects are summarized in Figure 2, where the upper part illustrates the responses of active capital (the dotted line) and investments (the solid line) when the capacity utilization rate is initially at 0.7, while the lower part illustrates the responses when the capacity utilization rate is initially at 0.95.

[Figure 2: Real options effect of uncertainty on investment when firms are able to change capacity utilization]

The result of the analysis on the obtained investment function $I_t = I \left(K_{t-1}, X_t, \sigma \right)$ can be summarized as follows. Firstly, the firm's actual use of capital (active capital) is rather flexible compared to the capital stock. The firm primarily responds to demand shocks by changing its capacity utilization. It decides to undertake new investments only when the capacity utilization rate has reached one. Thus, active capital directly responds even to demand shocks that are too small to motivate new investments in the real options AE model. Note that even responses of active capital has a range of inaction, as illustrated in Figure 2, simply because it only follows responses of investment when the capacity utilization has reached one. Note also that the slope of the active capital responses is steeper before the range of inaction than after. This is because, whereas active capital is determined by the first-order condition $\partial \pi \left(K_t^a, X_t \right) / \partial K_t^a = m$ as long as vacant capacity exists, new investment (and hence active capital) is determined by $\partial \pi \left(K_t, X_t \right) / \partial K_t = \left[m + (1 - \beta) q \right] \psi$ when the firm operates at

full capacity, where $\psi \, (> 1)$ denotes the real options effect. It is trivial to show that $m < [m + (1 - \beta) q] \psi$. If a production function is convex and the profit function $\pi (\cdot)$ is thereby also convex, it can be shown that active capital K_t^a $(< K_t)$ responds more than the capital stock K_t to a given change in demand shock X_t. Hence, the response of active capital to demand shocks may be approximated by a concave function in empirical analysis. Similarly, the response of capacity utilization to demand shocks may be approximated by a concave function, as capacity utilization is limited at one.

Secondly, the responsiveness of investment to demand shocks varies with current capacity utilization. For a given increase in output demand, the firm invests less when it has a larger stock of redundant capital. The responses are therefore a positive function of current capacity utilization. In contrast, the responses in active capital and capacity utilization are negative functions of the current capacity utilization. This is because the firm needs to undertake new investment to increase active capital when capacity utilization is already close to one.

Thirdly, the impacts of uncertainty on responsiveness of active capital and capacity utilization are unclear, since there are two different forces operating in these relationships. One is a "user cost" effect, analogous to the one referred to in Abel and Eberly (1999). The main argument of real options theory is that higher levels of uncertainty lead to a wider range of inaction for investment (the range [C, A] in Figure 2). When capacity utilization rate is already at a high level, the firm's cautious attitude toward new investments implies that the firm has less room to vary capacity utilization in response to demand shocks. The other effect is a "hangover" effect that is also analogous to the one discussed in Abel and Eberly (1999). This effect arises when large uncertainty generates a large negative demand shock. The firm faces redundant capital that it cannot sell off. If the firm initially is in this situation, it can easily respond to a positive demand shock by raising its capacity utilization. Therefore, the joint implication of the two effects for the responses of capacity utilization and active capital is unclear.

Fourthly, the cautious behavior, predicted by the real options AE model, diminishes when the firm has a possibility to adjust its capacity utilization. Under the assumption of full capacity utilization, the firm is unable to cut capital input in a recession. This means that it cannot cut as much labor

input as it could otherwise under the assumption of Cobb-Douglas technology.[8] In contrast, the firm under the assumption of variable capacity utilization can reduce its capital input, and hence also labor input. Therefore, the binding costs that new investments can imply for the firm in the presence of an uncertain future are lower, which makes the firm less cautious when deciding on new investments. This implies a shorter range of inaction in our model (the range [C, A] in Figure 2) than in the AE model (the range [0 A] in Figure 1).

In the empirical analysis, we test the hypotheses raised in the first and second points above. We also analyze the third point, i.e., the joint direction of the two different effects of uncertainty on the responses of active capital and capacity utilization. It is not possible to empirically test the fourth point.

3   Swedish sectoral data for manufacturing industries

3.1   Capacity utilization

Statistics Sweden published quarterly data on capacity utilization at the 2-digit sectoral level for the manufacturing industries during the period 1990–1998. Data was collected by surveying all firms with more than 200 employees and a selection of firms with 10–199 employees. The rate of capacity utilization was defined as the ratio between actual production and the currently available production capacity.[9]   The collected data was aggregated using firms' value added as weights.

Since 1999, however, Statistic Sweden publishes the capacity utilization rate only for the manufacturing industry as a whole and five selected 2-digit sectors (chemicals, basic metals, machinery, communication equipment, and motor vehicles). The number of employees is used as weights in aggregation.

---

[8] Under the assumption of Leontief technology, the firm keeps labor input constant since labor input is solely determined by current capital input.

[9] The question asked in the questionnaire is "How large was capacity utilization at the observation unit in the last quarter?" In the instructions, it is clarified that capacity utilization is defined as the ratio between actual production and the currently available production capacity (i.e., full capacity utilization). It is further explained that full capacity refers to the production level that can be achieved with existing machinery and current production methods, based on the working hours, shifts, and production mix that are considered to be normal at the plant for the quarter. Variations in the production capacity due to seasonal factors (e.g., holidays) shall not be counted in the reported rate of utilization. Although we note that this is a measure of the ratio of actual output to potential output instead of the ratio of active capital to the installed capital, we believe the former serves as a proxy for the latter.

The National Institute of Economic Research (NIER) in Sweden also publishes quarterly data on capacity utilization at the 2-digit sectoral level for the manufacturing industry, but only since 1996. Data is collected via the Economic Tendency Survey, which reaches all firms with more than 100 employees and a selection of the firms with 10–99 employees.[10] The collected data is aggregated using firms' value added as weights.

The advantage of the data from Statistics Sweden is that it starts in 1990. However, there are no sectoral data after 1998 except for the five sectors mentioned above. Thus, we complement the data from Statistics Sweden with data from NIER for the remaining sectors. Investigating the period 1996–1998, where both sources have data on capacity utilization, we find that they follow each other relatively well. However, there are level gaps up to 10 percentage points in some sectors. The utilization rate provided by NIER is often lower than in the data from Statistic Sweden. We suppose that the gaps reflect the differences in survey methods and sampling procedures between the two sources. To make a smooth time series on capacity utilization for each 2-digit sector, we proportionally shift the data from NIER by its gap to the data from Statistic Sweden. In converting quarterly to annual data, we simply use the annual average.

3.2 Uncertainty

The measures of uncertainty used in the present study is based on the Economic Tendency Survey by NIER. In their quarterly inquiry mentioned above, they ask about *production volume*, *selling prices*, and *order inflows* in the last three months. The response alternatives are "have increased," "have been unchanged," and "have decreased." In the same inquiry, they also ask the firms about their expectations on *production volume*, *selling prices*, and *order inflows* in the next three months, with three response alternatives: "will increase," "will remain unchanged," and "will decrease." NIER publishes only net figures of responses, i.e. the balance between the percentage of firms reporting an increase and those reporting a decrease. The net figures are available for each quarter at the 2-digit sectoral level since 1990. We denote the net figure for the answer to the former question $g(\Delta s_{j,t})$, assuming that the net figure is a monotonously increasing function $g(\cdot)$ of the change in a state variable (i.e., either *production volume*, *selling prices*, or *order inflows*) for sector j in the last

---

[10] The question asked in the questionnaire is "What is the estimated value of current capacity utilization, apart from normal seasonal changes?"

three months $\Delta s_{j,t}$. We denote the net figure for the answer to the latter question $h\left[E_t\left(\Delta s_{j,t+1}\right)\right]$, assuming that the net figure is a monotonously increasing function $h\left[\cdot\right]$ of the expected change in a state variable for sector j in the coming three months $E_t\left(\Delta s_{j,t+1}\right)$.

Based on these net figures, we construct our measure of uncertainty. We assume that the difference between the expectation and the outcome is an indication of uncertainty for the firms in the sector. Firstly, we take the difference between the net figure for the realization of a state variable and the net figure for the expectation for the state variable at one quarter earlier. Let us call the difference $z_{j,t}$, which is defined as

$$z_{j,t} = g(\Delta s_{j,t}) - h\left[E_{t-1}\left(\Delta s_{j,t}\right)\right]. \qquad (6)$$

We then construct a measure of uncertainty $\sigma_{j,t}$ for the sector j as

$$\sigma_{j,t} = \sum_{t_q=1}^{4} z_{j,t_q,t}^2 + \sum_{t_q=1}^{4} z_{j,t_q,t-1}^2, \qquad (7)$$

where t_q and t denote quarter and year, respectively. In other words, our uncertainty measures are based on the gap between the expectation and the outcome in the last two years. This is because we assume that it takes firms a few quarters to shape their perception of the current uncertainty in the sector. This is also to avoid multicollinearity with other explanatory variables in our empirical model, where we use one-year differences in output and other variables.

3.3   Other variables and sector classification

Data on value added, investment, and capital stock are obtained from the national accounts and the business survey of Statistics Sweden. All variables are deflated by sectoral annual price indices for the respective variables in 2000. The values of output and investment are available at the quarterly level, while the values of the capital stock is only available at the annual level. To convert quarterly data to annual data, we add all output and investments, respectively, within a year.

The variables are basically classified according to SNI2002, which is a Swedish Standard Industrial Classification (SNI) based on the EU's recommended standard NACE Rev.1.1. However, a new classification has been introduced since 2008, which is called SNI2007, based on the EU's NACE Rev.2. For most of

the variables we need, the authorities provide data according to both SNI2002 and SNI2007, even after 2008. In other cases, it is possible to reclassify the data according to SNI2002 without a large loss of precision. However, since some sectors are not as detailed in the new classification as in the old classification, we are unable to reclassify. Therefore, we have no data for those sectors in 2009 and 2010. Otherwise, all variables are available annually from 1990 to 2010.

3.4 Variation in capacity utilization for the manufacturing industry

Figure 3 describes the development of capacity utilization for the whole manufacturing industry and three subcategories (the basic metals sector, the machinery sector, and the motor vehicles sector[11]) during 1990–2010. As can be seen, the value fluctuates between 60 and 98 percent over the business cycle, implying that active capital (actual capital use) deviates considerably from the installed capital stock.

[Figure 3: Capacity utilization (1990–2010)]

Figures 4–7 present the annual changes in active capital and the capital stock together with the change in value added, for the whole manufacturing industry and the three subsectors. Active capital is defined as the capital stock multiplied with the capacity utilization rate:

$$K_{i,t}^a = K_{i,t} \cdot CU_{i,t}, \tag{8}$$

where $K_{i,t}^a$ is active capital, $K_{i,t}$ is the capital stock, and $CU_{i,t}$ is the capacity utilization rate. As shown, active capital fluctuates more than the capital stock, and also follows the changes in GDP better.

[Figures 4–7: The relationship between changes in capital and in output]

Figures 8–11 analyze the same issue from another viewpoint. The change in active capital is decomposed into the contributions of the change in capital stock (net investment) and of the change in capacity utilization, according to

---

[11] Theese are the sectors for which Statistics Sweden has data on capacity utilization for the whole sample period.

the approximation:

$$\Delta K_{i,t}^a \simeq \Delta K_{i,t} \cdot CU_{i,t} + K_{i,t} \cdot \Delta CU_{i,t}, \qquad (9)$$

where the first term denotes the contribution of the change in the capital stock and the second term denotes the contribution of the change in capacity utilization. It is shown in the figure that, for the whole manufacturing industry, changes in the capital stock (net investment) are the main contributor to the changes in active capital during the periods 1996–2002 and 2006–2007. However, the contribution of the change in capacity utilization is not negligible. In some years, the contributions of the two factors act in the opposite directions. The tendency that net investment and capacity utilization cancel each other out is also apparent in the three subsectors.

[Figures 8–11: Contributions of changes in the capital stock and capacity utilization to changes in active capital]

These figures highlight the problem that many firms face at every moment: Firms try to anticipate future demand shocks and make an investment plan to maximize the expected profits. However, because of a time lag before the installed capital is available and of uncertainty regarding future demand shocks, there is always a gap between the *ex post* desired capital stock and the actual capital stock – the gap that is filled by changes in capacity utilization. In this way, capacity utilization is an important component in the firm's investment behavior.

4   Empirical specification

4.1   The error correction model

The empirical model used in this study is based on the error correction model. Bloom et al.(2003) present an application of the error correction model to tests of the real options hypothesis. We assume that the desired capital stock in the long run is expressed by the equation

$$k_{i,t}^* = y_{i,t} + a_i + b_t, \qquad (10)$$

where $k_{i,t}^*$ is the log of the long-run optimal capital stock, $y_{i,t}$ is the log of output, a_i represents firm-specific factors such as a technical coefficient, a markup parameter, and depreciation, and b_t denotes time-specific factors such as trends in technical progress and demand shocks due to the business cycle.

In the short run, however, the actual capital stock deviates from the long-run optimal level because the firm is only partially able to adjust its capital stock due to some form of adjustment costs, such as irreversibility. Following Bloom et al.(2003), we express the short-run capital dynamics as

$$k_{i,t} = \alpha_1 k_{i,t-1} + \beta_0 y_{i,t} + \beta_1 y_{i,t-1} + a_i + b_t + \epsilon_{i,t}. \tag{11}$$

For transformation to an error correction form, we first subtract $k_{i,t-1}$ from both sides and then rearrange the right-hand side so that we have the components $\Delta y_{i,t}$, $y_{i,t-1}$ and $k_{i,t-1} - y_{i,t-1}$. The result is

$$
\begin{aligned}
\Delta k_{i,t} &= \beta_0 \Delta y_{i,t} + [(\beta_0 + \beta_1) + (\alpha_1 - 1)] y_{i,t-1} \\
&\quad + (\alpha_1 - 1)(k_{i,t-1} - y_{i,t-1}) + a_i + b_t + \epsilon_{i,t}.
\end{aligned} \tag{12}
$$

For simplicity, we rewrite the equation as

$$\Delta k_{i,t} = \beta_0 \Delta y_{i,t} + \gamma y_{i,t-1} + \theta (k_{i,t-1} - y_{i,t-1}) + a_i + b_t + \epsilon_{i,t}, \tag{13}$$

where $\gamma = (\beta_0 + \beta_1) + (\alpha_1 - 1)$ and $\theta = \alpha_1 - 1$. Since $\Delta k_{i,t}$ can be interpreted as an approximation of the net investment rate, i.e.,

$$
\begin{aligned}
\Delta k_{i,t} &\simeq \frac{K_{i,t} - K_{i,t-1}}{K_{i,t-1}} \\
&= \frac{I_{i,t-1}}{K_{i,t-1}} - \delta_{i,t-1},
\end{aligned} \tag{14}
$$

where $\delta_{i,t}$ is the depreciation rate, Eq.(13) represents a dynamic specification of the investment function.[12] Since we assume that investment is a positive function of an output change, the coefficient β_0 is expected to be positive. The coefficient θ on the error-correction term is expected to be negative, so that the capital stock will be adjusted upwards when it is below the long-run optimal level and vice versa. The coefficient γ on the lagged output reflects the long-run elasticity of capital with respect to output. When production technology is

---

[12]We assume that the depreciation rate $\delta_{i,t}$ is different across sectors but constant over time, so that it will be captured by the time-invariant sector dummy a_i.

characterized by constant returns to scale, γ is equal to zero;[13] when goods are produced under decreasing returns to scale, it is negative.[14]

4.2 Responsiveness of the capital stock and capacity utilization

To empirically test the hypotheses presented in Section 2, we modify the original error correction model in Eq.(13) with additional terms:

$$
\begin{aligned}
\Delta k_{i,t}^a &= \beta_0 \Delta y_{i,t} + \eta \left(\Delta y_{i,t}\right)^2 + \gamma y_{i,t-1} + \theta \left(k_{i,t-1}^a - y_{i,t-1}\right) + \lambda cu_{i,t-1} \\
&\quad + \phi_1 \sigma_{i,t-1} + \phi_2 \sigma_{i,t-1} \Delta y_{i,t} + a_i + b_t + \epsilon_{i,t},
\end{aligned}
\tag{15}
$$

where $k_{i,t}^a$ is the log of active capital $K_{i,t}^a$ (see Eq.(8)) and $cu_{i,t-1}$ is the log of $CU_{i,t-1}$. We add $cu_{i,t-1}$ because we expect the level of capacity utilization at the moment of decision making (i.e., at $t-1$) to affect responses of capital use between $t-1$ and t. We add the square of $\Delta y_{i,t}$ to capture convexity or concavity of the function as in Bloom et al.(2003); see Eq.(1). The terms $\phi_1 \sigma_{i,t-1}$ and $\phi_2 \sigma_{i,t-1} \Delta y_{i,t}$ capture the effects of uncertainty $\sigma_{i,t-1}$, which is also dated at $t-1$. Note also that we have $k_{i,t-1}^a$ on the right-hand side simply to follow the dependent variable.

Since $k_{i,t}^a$ is expressed by the sum of $k_{i,t}$ and $cu_{i,t}$ (see Eq.(8)), Eq.(15) can be divided into two separate equations:

$$
\begin{aligned}
\Delta k_{i,t} &= \beta_0' \Delta y_{i,t} + \eta' \left(\Delta y_{i,t}\right)^2 + \gamma' y_{i,t-1} + \theta' \left(k_{i,t-1}^a - y_{i,t-1}\right) + \lambda' cu_{i,t-1} \\
&\quad + \phi_1' \sigma_{i,t-1} + \phi_2' \sigma_{i,t-1} \Delta y_{i,t} + a_i' + b_t' + \epsilon_{i,t}',
\end{aligned}
\tag{16}
$$

and

$$
\begin{aligned}
\Delta cu_{i,t} &= \beta_0'' \Delta y_{i,t} + \eta'' \left(\Delta y_{i,t}\right)^2 + \gamma'' y_{i,t-1} + \theta'' \left(k_{i,t-1}^a - y_{i,t-1}\right) + \lambda'' cu_{i,t-1} \\
&\quad + \phi_1'' \sigma_{i,t-1} + \phi_2'' \sigma_{i,t-1} \Delta y_{i,t} + a_i'' + b_t'' + \epsilon_{i,t}''.
\end{aligned}
\tag{17}
$$

Eq.(16) explains the change in the capital stock by demand shifts and other terms. We expect the effect of demand shifts, β_0', to be positive and the co-

---

[13] This can be seen by substituting $k_{i,t}$ and $k_{i,t-1}$, and $y_{i,t}$ and $y_{i,t-1}$ in Eq.(11) with the long-run optimal values $k_{i,t}^*$ and $y_{i,t}^*$, respectively. The steady-state relationship between $k_{i,t}^*$ and $y_{i,t}^*$ is given by $y_{i,t}^* = (1-\alpha)/(\beta_0 + \beta_1) k_{i,t}^*$. The constant returns to scale is equivalent to $(1-\alpha)/(\beta_0 + \beta_1) = 1$, or $\gamma = 0$.

[14] In Bloom et al.(2003), constant returns to scale are assumed.

efficient on the error correction term, θ', to be negative. λ' is expected to be positive because, when capacity utilization is already at a higher level, the firm has less room to adjust its active capital by raising its capacity utilization for a given increase in output demand, and the firm therefore needs new investment. In the opposite case, i.e., when capacity utilization is lower, the firm is less willing to invest. η' captures the convexity or concavity of the function with respect to $\Delta y_{i,t}$, and we expect it to be positive as in Bloom et al.(2003). ϕ'_2 is a core coefficient to test the real options hypothesis, and is expected to be negative.

Eq.(17) explains the change in capacity utilization. While we expect β''_0 to be positive, expectations on the signs of other coefficients are different from those in Eq.(16). η'' is expected to be negative since the ceiling of the capacity utilization rate at one leads to diminishing adjustment as demand shocks become large. The error correction term has no practical implication in this equation. λ'' is expected to be negative because, when capacity utilization is already at a higher level, the firm has less room to raise capacity utilization for a given increase in output demand. ϕ''_2 is a core coefficient for the test of whether the firm tends to use internal adjustment by changing capacity utilization rather than net investment at higher uncertainty levels. If this hypothesis is true, ϕ''_2 is positive.

5 Results for simulated data

The structural model with variable capacity utilization proposed in Chapter 1 is difficult to test empirically at the firm level for a couple of reasons. Firstly, although we suggest the empirical Eqs.(15)–(17) to test the model, it is not possible to tell whether the signs of the estimated coefficients based on data generated from the structural model are as we expect. Secondly, data on capacity utilization is not available at the firm level but only at the sectoral level. A question arises if the mechanism for investment decisions and capital use decisions that operates at the firm level can be detected in econometric analyses using sectoral data.

Therefore, first we let the structural model generate series of data at the firm level and analyze whether the signs of the coefficients estimated from the empirical Eqs.(15)–(17) are as we expect. Then we aggregate the firm-level data to the sectoral level and estimate the same set of the empirical equations to an-

alyze whether we can detect the parameters we needed to verify our hypotheses at the sectoral level.

To generate simulated data, we first need to obtain the policy function for investment $I_t = I(K_{t-1}, X_t, \sigma)$ through numerical iteration. We assume that the standard deviation σ varies over time, and denote it σ_t. We further assume that σ_t evolves as persistent stationary AR(1) process $\sigma_{t+1} = \rho(\sigma_t - \mu) + \mu + \epsilon_t$, where μ is the mean of σ_t and $\epsilon_t = iid.N(0, s_\epsilon)$. For simplicity, we discretize σ_t so that it only takes five different values $[0.10, 0.15, 0.20, 0.25, 0.30]$. The dynamic optimization problem, Eq.(2), is modified as follows:

$$V(K_{t-1}, X_t, \sigma_t) = \max_{\substack{I_t \geq 0 \\ K_t^a \leq K_t}} \pi(K_t^a, X_t) - mK_t^a - qI_t + \beta E_t[V(K_t, X_{t+1}, \sigma_{t+1})].$$

(18)

Transition from σ_t to σ_{t+1} is defined by a 5×5 transition matrix. We use the algorithm proposed by Tauchen (1986) to discretize an AR(1) process. We set the autoregressive coefficient ρ to 0.80 to reflect the fact that uncertainty is rather persistent. We also set $s_\epsilon = 0.03$ so that the standard deviation of σ_t over time is equal to $s_\sigma = s_\epsilon/\sqrt{1 - \rho^2} = 0.05$. We further set the upper and lower bounds of the range as twice s_σ from the mean 0.20, so that the five points on the grid are equal to the values defined for the discretization of σ_t.

We assign the following values to the parameters in the dynamic optimization problem presented by Eq.(18), the underlining demand function $Q_t = X_t P_t^{-\epsilon}$, the production function $Q_t = L_t^{1-\alpha} K_t^\alpha$, and the geometric Brownian motion presented by Eq.(3): $\alpha = 0.33$, $\mu = 0.029$, $\beta = 0.9523 (= 1/(1+r)$ where $r = 0.05$), and $\epsilon = 10$. These are the same values as used in Abel and Eberly (1999). The ratio of m to q is set to 0.25, which generates moderate fluctuation in capacity utilization, as shown in Chapter 1. All the other parameters can be chosen arbitrarily as they only affect the levels of the capital stock and active capital without influencing the ratio between them.

The obtained policy function $I_t = I(K_{t-1}, X_t, \sigma_t)$ is applied to artificially generated demand shocks at the firm level that follow a geometric Brownian motion:

$$\ln X_{i,t} \equiv x_{i,t} = x_{i,t-1} + \tilde{\mu}_{i,t} + \sigma_{i,t} u_{i,t},$$

(19)

where $\tilde{\mu}_{i,t}$ is a growth rate that is defined as $\tilde{\mu}_{i,t} = \mu - 0.5\sigma_{i,t}^2$ and $\sigma_{i,t} u_{i,t}$ is a firm-specific component of uncertainty with random variable $u_{i,t} = iid.N(0, 1)$.

The subscript i indicates that the demand shock is firm specific. To create enough variation in aggregated uncertainty at the sectoral level, we assume that $\sigma_{i,t}$ is sector specific and applies to all firms within a sector. If $\sigma_{i,t}$ is also firm-specific, the aggregated uncertainty at the sectoral level may be similar across sectors. The initial value of the sector-specific $\sigma_{i,t}$ is randomly chosen from the values [0.10, 0.15, 0.20, 0.25, 0.30]. $\sigma_{i,t}$ evolves according to the $5{\times}5$ transition matrix defined in Tauchen (1986) with the parameters described above.

The initial value of the demand shock, $X_{i,0}$, is set to 0 for all i. The initial value for the capital stock is set to half of the optimal level at $X_{i,0} = 0$ under the assumption of no irreversibility constraint.

The number of sectors is set to 20 (close to our empirical dataset where it is 19) and we assume that each sector has 500 firms. With this setting, we let the structural model generate series of data for 40 years and exclude the first 20 years. This is to avoid influence of the initial values. The sult from the simulation shows that it takes around 20 years for the capacity utilization rate to stabilize.

The generated firm-level data is then aggregated to the sector level. For all variables except capacity utilization, the firm-level data is summed and divided by the number of the firms in order to avoid the figures becoming too large. For sector-aggregated capacity utilization, we take a weighted average of the firm-level values within a sector using the firms' outputs as weights.

Tables 1 and 2 show the descriptive statistics for the firm-level data and the sector-aggregated data, respectively.

[Table 1: Descriptive statistics for the simulated firm-level data]

[Table 2: Descriptive statistics for the sector-aggregated data]

We will now estimate the empirical equations. Firstly, we investigate the signs of the coefficients in Eqs.(15), (16), and (17) when the equations are estimated using the firm-level data. In the regression of each equation, we add dummy variables for years and sectors to run fixed-effect (FE) panel estimation. These models potentially suffer from biases. Taking Eq.(15) as an example, the demeaned regressor $\left(k_{i,t-1}^a - y_{i,t-1}\right) - (T-1)^{-1} \sum_1^{T-1} \left(k_{i,s}^a - y_{i,s}\right)$ is negatively correlated with the demeaned residual $\epsilon_{i,t} - (T-1)^{-1} \sum_1^{T-1} \epsilon_{i,t}$ because $\epsilon_{i,t-1}$ is part of the process that generates $k_{i,t-1}^a$. The bias is, if anything, negative. However, it moves toward zero as T moves toward infinity. Even in Eqs.(16)

and (17), there is similar endogenuity between the demeaned residual and some of the demeaned regressors.

To tackle the bias, we estimate Eqs.(15), (16), and (17) with the instrumental variable method in a fixed-effect model, using lagged regressors as instruments. Strictly speaking, lagged regressors are not valid instruments in a fixed-effect model as they are correlated with the demeaned error term. A first-difference model with the instrumental variable method is therefore preferred. However, Since we have a fairly long time-series ($T = 20$), the resulting bias is expected to be small.

Table 3 presents our empirical results using fixed effect with lagged regressors as instruments (FE-IV). In these estimations, we use the following instruments: $k_{i,t-2}$ and $cu_{i,t-2}$ for $k^a_{i,t-1} - y_{i,t-1}$ and $cu_{i,t-1}$ in Eq.(15), $k_{i,t-2}$ for $k^a_{i,t-1} - y_{i,t-1}$ in Eq. (16), and $cu_{i,t-2}$ for $cu_{i,t-1}$ in Eq.(17). From the first column where the estimation of Eq.(16) is presented, we see that $\Delta y_{i,t}$ has a positive effect on investment (i.e., $\Delta k_{i,t}$) and that the effect is convex. The error-correction term $k^a_{i,t-1} - y_{i,t-1}$ has a negative coefficient. Current capacity utilization has a positive impact on the investment decision. The negative coefficient on the cross term $\sigma_{i,t-1} \Delta y_{i,t}$ indicates that the firm is more cautious regarding investments at higher levels of uncertainty. These results are consistent with our expectation.

[Table 3: Regression analysis using the simulated firm-level data]

From the second column where the estimation of Eq.(17) is presented, we see that the coefficient on $\Delta y_{i,t}$ is significantly positive. The effect on changes in capacity utilization is concave. Current capacity utilization $cu_{i,t-1}$ has a negative impact. These findings are also in line with our expectation. The positive coefficient on the cross term $\sigma_{i,t-1} \Delta y_{i,t}$ indicates that the firm is likely to adjust capacity utilization more when the uncertainty level is higher.

The third column shows the estimation of Eq.(15). Although this is estimated separately from the other two equations, each coefficient is close to the sum of the corresponding coefficients in the two equations. It is observed that changes in active capital is a positive and slightly concave function of demand shifts $\Delta y_{i,t}$. The positive coefficient on $\sigma_{i,t-1} \Delta y_{i,t}$ indicates that the firm adjusts its active capital more for a given demand shock at higher uncertainty levels.

Since the instrumental variable method in a FE model suffers from a bias, we also estimate the models using the dif-GMM model proposed by Arellano and

Bond (1991). The result is presented in Table 4. For each of the three equations, we first estimate the dif-GMM model with the second lag of the regressors specified in the "instruments" row as instrument variables. The specification tests such as the AR(2) and the Sargan tests indicate specification errors in the model in all three cases. We therefore estimate the same model with the third lag as instruments, yet the problem remains except for the $\Delta k_{i,t}^a$-equation where the estimated coefficients are quite similar to those estimated with FE-IV.[15]

[Table 4: Regression analysis using the simulated firm-level data with dif-GMM]

We now investigate if we can obtain similar regression results when we estimate the three equations using the sector-aggregated simulated data. Table 5 presents the estimation results using the fixed-effect method with instrument variables (FE-IV). Table 6 shows the estimation results using dif-GMM as well as sys-GMM as proposed by Arellano and Bover (1995) and Blundell and Bond (1998). In Table 6, it is observed that we do not reject the null hypothesis of no specification errors at the 5 percent significance level in all dif-GMM estimations. However, we reject the null hypothesis in the sys-GMM estimations except for the $\Delta cu_{i,t}$-equation. We therefore focus on the results with dif-GMM.

[Table 5: Regression analysis using the simulated sector-aggregated data]

[Table 6: Regression analysis using the simulated sector-aggregated data with dif-GMM and sys-GMM]

Compared to Table 4, most of the coefficients have the same signs and similar significance. There are, however, some exceptions. The demand shift has no significant positive effect in the $\Delta cu_{i,t}$-equation. The squared term $(\Delta y_{i,t})^2$ is only significant in the first equation. The sign on the error-correction term $k_{i,t-1} - y_{i,t-1}$ is the opposite in the $\Delta cu_{i,t}$-equation, but the coefficient has no practical interpretation. The lagged capacity utilization $cu_{i,t-1}$ and the cross term $\sigma_{i,t-1}\Delta y_{i,t}$ have no significant effects in the $\Delta k_{i,t}^a$-equation.

---

[15]Another reason for potential biases is that there may be feedback from the dependent variable $k_{i,t-1}$ to $y_{i,t}$ on the right-hand side. $y_{i,t}$ is in this case endogenous and is therefore correlated with the error term $\epsilon_{i,t}$. We have therefore also estimated the empirical equations using the FE-IV and dif-GMM method where we additionally instrument $\Delta y_{i,t}$, $(\Delta y_{i,t})^2$ and $\sigma_{i,t-1}\Delta y_{i,t}$ with $y_{i,t-2}$, $(\Delta y_{i,t-1})^2$ and $\sigma_{i,t-1}y_{i,t-2}$. However, the result for the FE-IV method shows that the standard errors are too large and nearly all coefficients are insignificant. The result for the dif-GMM shows that the specification tests are unsatisfactory in the sense that they reject the null of no specification errors.

To sum up this section, the analysis using the simulated data shows that the signs of the estimated coefficients at the firm level are as we have expected, and that the mechanism for the decisions about investment and capital use at the firm level can be detected by empirical estimations at the sectoral level.

6   Results for real data

We implement the same analysis with the real data described previously using the FE-IV and the GMM methods. We have sectoral data for 19 sectors in Swedish manufacturing industries during the period 1990-2010 except for some sectors that lack data for 2009 and 2010.

Table 7 shows the descriptive statistics for the data. Table 8 presents our empirical results using the fixed-effect method (FE-IV) with the following instruments: $k_{i,t-2}$ and $cu_{i,t-2}$ for $k_{i,t-1}^a - y_{i,t-1}$ and $cu_{i,t-1}$ in Eq.(15), $k_{i,t-2}$ for $k_{i,t-1}^a - y_{i,t-1}$ in Eq.(16), and $cu_{i,t-2}$ for $cu_{i,t-1}$ in Eq.(17).

[Table 7: Descriptive statistics for the real data]

[Table 8: Regression analysis using the real data]

Columns 1–3 show estimation results of Eqs.(15), (16), and (17), where we use gaps between the expectation and the outcome of *production volume* to construct a measure of uncertainty. The result is consistent with what we have expected and also similar to the result from the simulated data. From Column 1, we see that $\Delta y_{i,t}$ has a positive and convex effect on investment (i.e., $\Delta k_{i,t}$). The error correction term is negative. Current capacity utilization $cu_{i,t-1}$ has a positive impact. The negative coefficient on $\sigma_{i,t-1}\Delta y_{i,t}$ indicates that the firm becomes more cautious at higher levels of uncertainty.

In Column 2, $\Delta y_{i,t}$ has a positive but concave effect on changes in capacity utilization. Current capacity utilization has a negative impact. The positive coefficient of $\sigma_{i,t-1}\Delta y_{i,t}$ indicates that the firm is likely to adjust capital use more by changing capacity utilization for a given demand shock when the uncertainty level is higher.

Column 3 indicates that changes in active capital is an increasing and concave function of demand shocks $\Delta y_{i,t}$.

The bottom part of the table presents the marginal effects of $\Delta y_{i,t}$ at different levels of uncertainty (the 10 percentage point, median, the 90 percentage point),

evaluated at the average $\Delta y_{i,t}$ (= 0.031). As shown, while the marginal effects on changes in active capital $\Delta k_{i,t}^a$ are partly explained by the effects through changes in the capital stock $\Delta k_{i,t}$ at a lower level of uncertainty, changes in active capital $\Delta k_{i,t}^a$ are fully explained by changes in capacity utilization $\Delta cu_{i,t}$ at a higher level of uncertainty. This result is consistent with the negative coefficient on $cu_{i,t-1}$ in Column 3, indicating that adjustments in active capital become slow when the capacity utilization is already at a high level as active capital is mainly adjusted through changes in capacity utilization.

Columns 4–6 exhibit estimation results where we use gaps between the expectation and the outcome of *order inflows* to construct a measure of uncertainty. The results are quite similar to Columns 1–3 with some differences. The effect of demand shifts $\Delta y_{i,t}$ on $\Delta k_{i,t}$ is convex, yet its marginal effects are, on average, not significantly different from zero. Furthermore, uncertainty now has a positive effect on $\Delta k_{i,t}$. However, since the average marginal effect is insignificant at all values of σ, uncertainty has no quantitatively significant impact on investment. The effect of uncertainty on capacity utilization is insignificant. The positive and significant coefficient on the cross term $\sigma_{i,t-1}\Delta y_{i,t}$ in the $\Delta k_{i,t}^a$-equation indicates that active capital responds more to demand shifts at higher levels of uncertainty.

Columns 7–9 present estimation results where gaps between the expectation and the outcome of *selling prices* are used to construct a measure of uncertainty. Again, the results are quite similar to Columns 1–3 with some differences. For example, although the cross term $\sigma_{i,t-1}\Delta y_{i,t}$ does not have a significant impact on investment, the average marginal effect of $\Delta y_{i,t}$ is slightly significant at a low level of uncertainty, but the positive effect disappears at higher levels of uncertainty.

As the instrumental variable method in an FE model suffers from a bias, we also estimate the models with dif-GMM. The results are presented in Table 9. When the specification tests for the estimations using the second lag as instruments are satisfactory, we present the estimated coefficients. Otherwise we present the coefficients estimated using the third lag as instruments. Table 9 shows that only the $\Delta cu_{i,t}$-equation satisfies the specification tests in the sense that they do not reject the null hypothesis of no specification errors. The estimated coefficients are similar to those in the FE-IV estimation.[16]

---

[16] We have also estimated the empirical equations using the sys-GMM method. All of them, however, rejected the null hypothesis of no specification errors.

[Table 9: Regression analysis using the real data with dif-GMM]

The findings from these analyses can be summarized as follows. Firstly, active capital adjusts more to demand shifts than does the capital stock. Since most of this adjustment goes through changes in capacity utilization and the capacity utilization is limited at one, the adjustment speed diminishes with the size of demand shocks $\Delta y_{i,t}$, which is reflected in the negative coefficient on $(\Delta y_{i,t})^2$ in the $\Delta k_{i,t}^a$-equation. In contrast, the effects of demand shocks on investment (i.e., changes in the capital stock, $\Delta k_{i,t}$) are convex, yet the effects are relatively small.

Secondly, because capacity utilization has an upper limit of one, the current level of capacity utilization plays a significant role in the changes in the capital stock and capacity utilization.

Thirdly, the effect of uncertainty on the responsiveness of investment ($\Delta k_{i,t}$) to demand shocks $\Delta y_{i,t}$ tends to be negative, which is in line with Bloom et al.(2003). On the other hand, the responsiveness of the active capital $\Delta k_{i,t}^a$ and capacity utilization $\Delta cu_{i,t}$ to demand shocks $\Delta y_{i,t}$ tend to be positively affected by uncertainty.

7 Conclusion

This study analyzes the impacts of uncertainty on investment and production dynamics. The existing real-options model with full capacity utilization, such as in Abel and Eberly (1999) and Caballero (1999), predicts that the responsiveness of investments to demand shocks is weaker when firms face higher levels of uncertainty. This implies that investment responses to policy stimuli may be seriously limited at higher levels of uncertainty, for example immediately after an uncertainty shock.

We generalize the Abel-Eberly model so as to accommodate fluctuations in capacity utilization, and analyze the effects of uncertainty in this modified model. In our model, capital adjustment is achieved in two ways: adjustment of the capital stock (i.e., investment) and adjustment of actual use of the capital stock, which we call "active capital".

Using both simulated data and sectoral data for the Swedish manufacturing industry for the period 1990-2010, we find that the responses of the capital stock to demand shocks tend to be weaker at a higher level of uncertainty. The results also show that the responses are approximated by a convex function of demand

shocks. These findings are consistent with the cautious behavior predicted by the real options model with full capacity, implying that the effects of uncertainty are not quantitatively different when allowing for variable capacity utilization.

However, we also find that the cautiousness disappears when we analyze the responses of active capital to demand shocks. Rather, the results indicate that active capital responds more to demand shocks at a higher level of uncertainty. This can be explained in two ways: (i) firms prefer internal adjustment through changes in capacity utilization to adjusting the capital stock at a high level of uncertainty, and (ii) when uncertainty is large, firms tend to sit with vacant capacity (a so-called "hangover" effect) and can easily adjust active capital in response to positive demand shocks. We have also found that the responsiveness of capacity utilization and active capital to demand shocks is approximated by concave functions.

The responses of the capital stock to demand shocks positively depend on current capacity utilization. In contrast, the responses of active capital and capacity utilization negatively depend on current capacity utilization. These findings are in line with the predictions from the structural model. Hence, a firm's current capacity utilization plays a significant role in the investment decision.

These results imply that the adjustment of active capital is rather flexible even at high levels of uncertainty. While the main argument of the real options AE model with full capacity utilization still holds, i.e. that short-run effects of policy stimuli on investment are limited in the immediate aftermath of an uncertainty shock, we argue that policy stimuli such as monetary or fiscal policy have positive impacts on production activities even in such situations.

References

[1] Abel, Andrew B. (1983), "Optimal Investment under Uncertainty," *American Economic Review*, 73(1), 228-33.

[2] Abel, Andrew B. and Eberly, Janice C. (1994), "A Unified Model of Investment under Uncertainty," *American Economic Review*, 84(5), 1369-84.

[3] Abel, Andrew B. and Eberly, Janice C. (1996), "Optimal investment with costly reversibility," *Review of Economic Studies*, 63(4), 581-93.

[4] Abel, Andrew B. and Eberly, Janice C. (1999), "The effects of irreversibility and uncertainty on capital accumulation," *Journal of Monetary Economics*, 44(3), 339-77.

[5] Arellano, Manuel and Bond, Stephen (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277-97.

[6] Arellano, Manuel and Bover, Olympia (1995), "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68(1), 29-51.

[7] Arrow, Kenneth J. (1968), "Optimal Capital Policy with Irreversible Investment," in Wolfe, James N. (ed.), *Value, Capital and Growth. Papers in Honour of Sir John Hicks*, Edinburgh University Press.

[8] Bernanke, Ben S. (1983), "Irreversibility, Uncertainty, and Cyclical Investment," *The Quarterly Journal of Economics*, 98(1), 85-106.

[9] Bertola, Guiseppe and Caballero, Ricardo J. (1994), "Irreversibility and Aggregate Investment," *Review of Economic Studies*, 61(2), 223-46.

[10] Bigsten, Arne; Collier, Paul; Dercon, Stefan; Fafchamps, Marcel; Gauthier, Bernard; Gunning, Jan W.; Oostendorp, Remco; Pattillo, Catherine; Söderbom, Måns; and Teal, Francis (2005), "Adjustment Costs and Irreversibility as Determinants of Investment: Evidence from African Manufacturing," *The B.E. Journal of Economic Analysis & Policy*, 4(1), 12.

[11] Bloom, Nicholas (2000), "The dynamic effects of real options and irreversibility on investment and labour demand," IFS Working Papers W00/15, Institute for Fiscal Studies.

[12] Bloom, Nicholas (2007), "Uncertainty and the Dynamics of R&D," *American Economic Review*, 97(2), 250-55.

[13] Bloom, Nicholas (2009), "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623-85.

[14] Bloom, Nicholas; Bond, Stephen R.; and van Reenen, John (2003), "Uncertainty and Company Investment Dynamics: Empirical Evidence for UK Firms," CEPR Discussion Papers 4025.

[15] Bloom, Nicholas; Bond, Stephen R.; and van Reenen, John (2007), "Uncertainty and Investment Dynamics," *Review of Economic Studies*, 74(2), 391-415.

[16] Blundell, Richard and Bond, Stephen (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115-43.

[17] Bond, Stephen R.; Söderbom, Måns; and Wu, Guiying (2011), "Pursuing the wrong options? Adjustment costs and the relationship between uncertainty and capital accumulation," *Economics Letters*, 111(3), 249-51.

[18] Brennan, Michael J. and Schwartz, Eduardo S. (1985), "Evaluating Natural Resource Investments," *The Journal of Business*, 58(2), 135-57.

[19] Caballero, Ricardo J. (1991), "On the Sign of the Investment-Uncertainty Relationship," *American Economic Review*, 81(1), 279-88.

[20] Caballero, Ricardo J. (1999), "Aggregate Investment," in Taylor, J. and Woodford, M. (eds.) *Handbook of Macroeconomics*, Amsterdam: North-Holland, 813–62.

[21] Dixit, Avinash K. (1989), "Entry and Exit Decisions under Uncertainty," *Journal of Political Economy*, 97(3), 620-38.

[22] Dixit, Avinash K. and Pindyck, Robert S. (1994), *Investment Under Uncertainty*, Princeton University Press, Princeton.

[23] Hartman, Richard (1972), "The effects of price and cost uncertainty on investment," *Journal of Economic Theory*, 5(2), 258-66.

[24] Ingersoll, Jonathan E. Jr. and Ross, Stephen A. (1992), "Waiting to Invest: Investment and Uncertainty," *Journal of Business*, 65(1), 1-29.

[25] Lucas, Robert E. Jr. and Prescott, Edward C. (1971), "Investment under Uncertainty," *Econometrica*, 39(5), 659-81.

[26] Majd, Saman and Pindyck, Robert S. (1987), "Time to build, option value, and investment decisions," *Journal of Financial Economics*, 18(1), 7-27.

[27] McDonald, Robert and Siegel, Daniel (1986), "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101(4), 707-28.

[28] Nickell, Stephen J. (1974), "On the Role of Expectations in the Pure Theory of Investment," *Review of Economic Studies*, 41(1), 1-19.

[29] Pindyck, Robert S. (1982), "Adjustment Costs, Uncertainty, and the Behavior of the Firm," *American Economic Review*, 72(3), 415-27.

[30] Pindyck, Robert S. (1988), "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review*, 78(5), 969-85.

[31] Tauchen, George (1986), "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177-81.
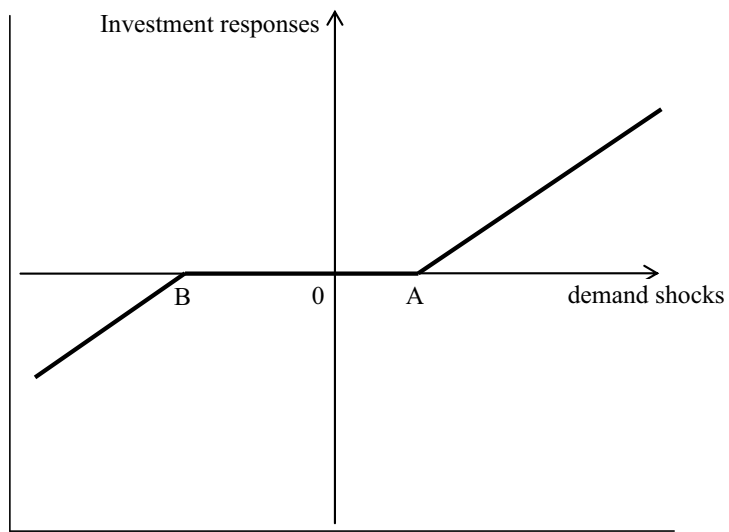
[Figure 1: Real options effect of uncertainty on investment]
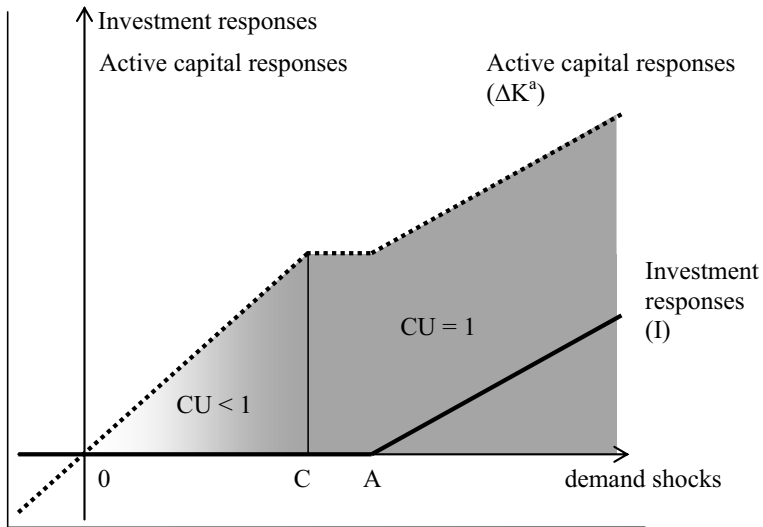
[Case 1] the capacity utilization rate is initially at 0.7
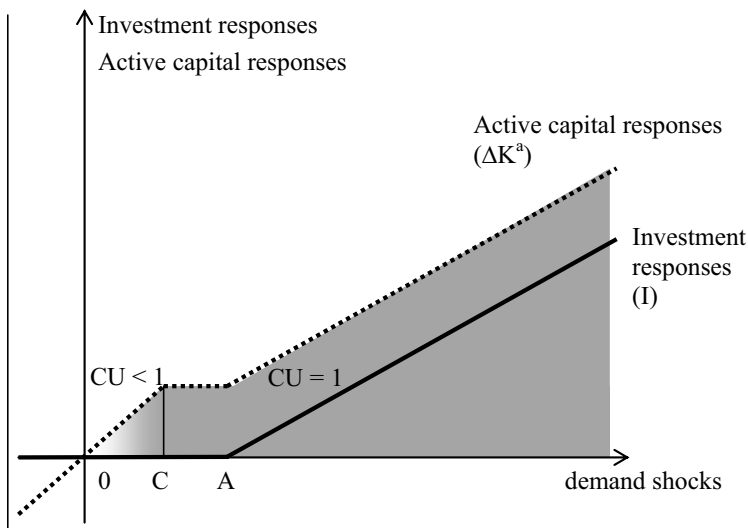
[Case 2] the capacity utilization rate is initially at 0.95

[Figure 2: Real options effect of uncertainty on investment when firms are able to change capacity utilization]

Capacity utilization
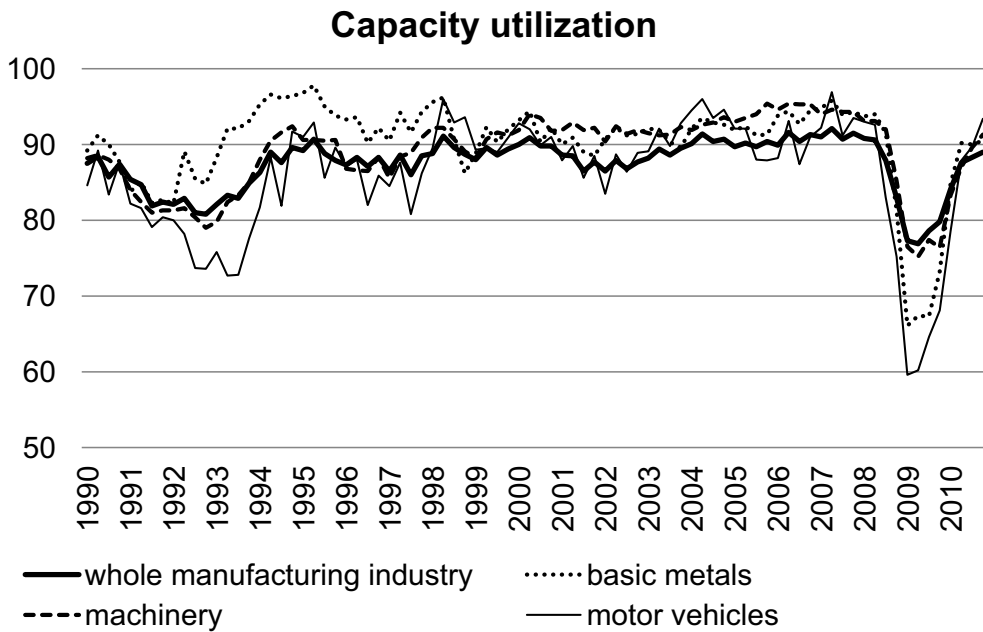
[Figure 3: Capacity utilization (1990-2010)]

Manufacturing industry
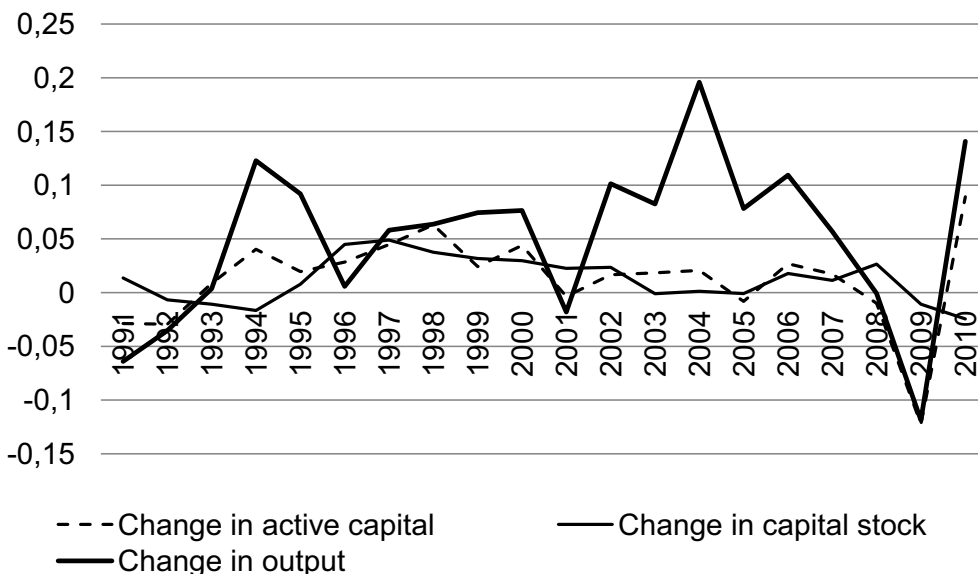
[Figure 4: The relationship between changes in capital and in output]

Basic metals
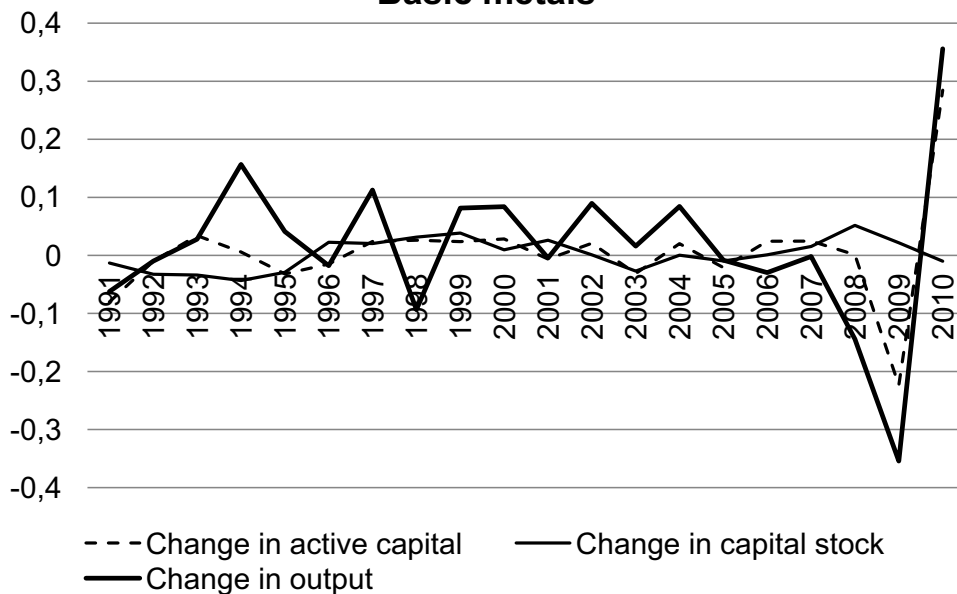
[Figure 5: The relationship between changes in capital and in output (cont.)]

Machinery

- - - Change in active capital ——— Change in capital stock
——— Change in output
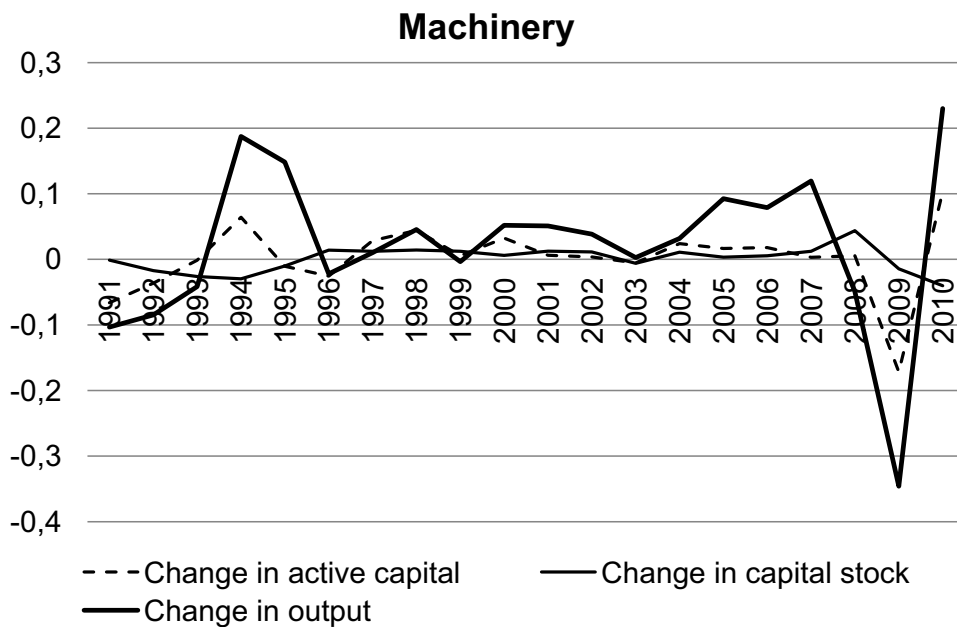
[Figure 6: The relationship between changes in capital and in output (cont.)]

Motor vehicles

- - - Change in active capital ——— Change in capital stock
——— Change in output
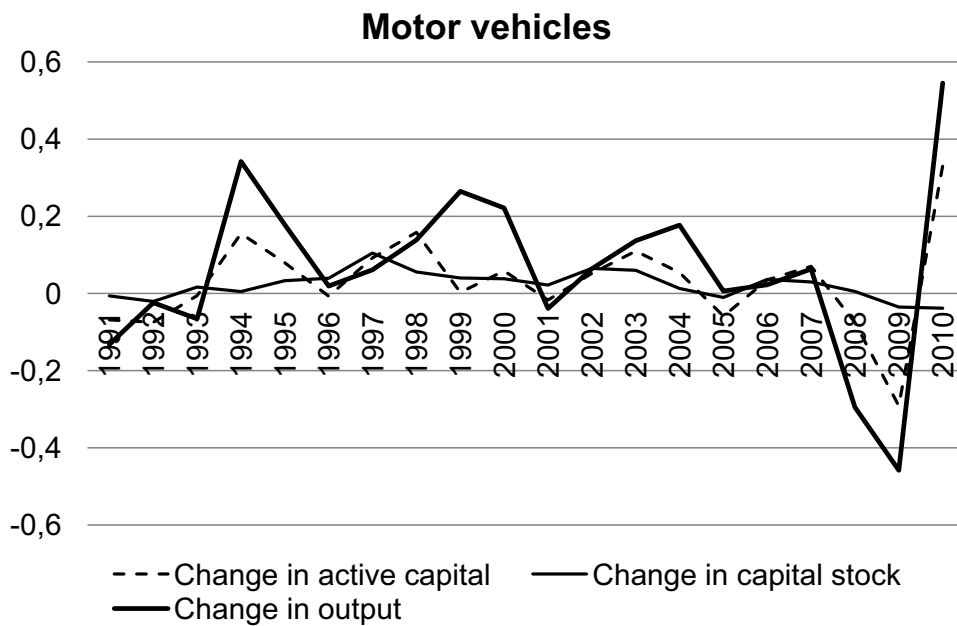
[Figure 7: The relationship between changes in capital and in output (cont.)]

Manufacturing industry

■ Contribution of the change in the capital stock
□ Contribution of the change in capacity utilization
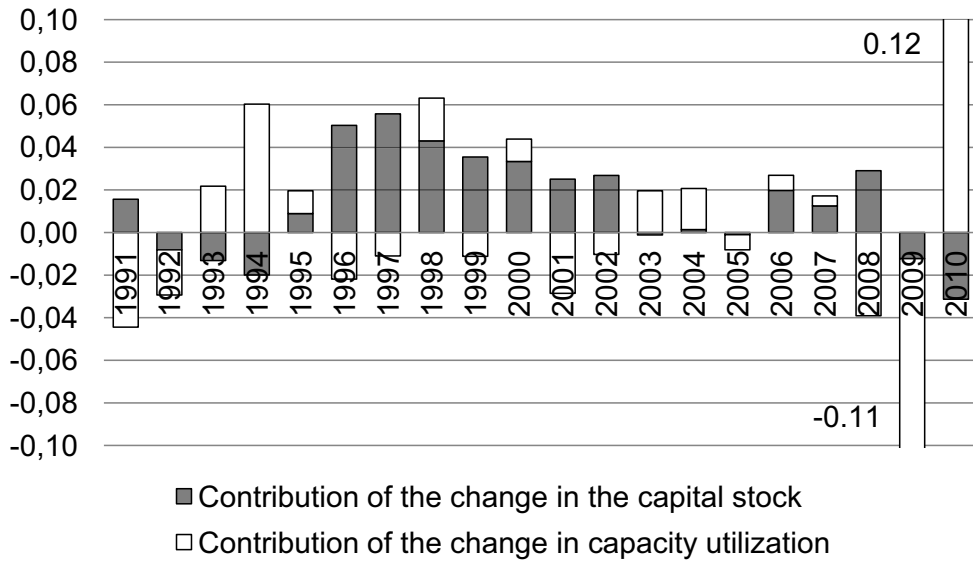
[Figure 8: Contributions of changes in the capital stock and capacity utilization to changes in active capital]

Basic metals

■ Contribution of the change in the capital stock
□ Contribution of the change in capacity utilization
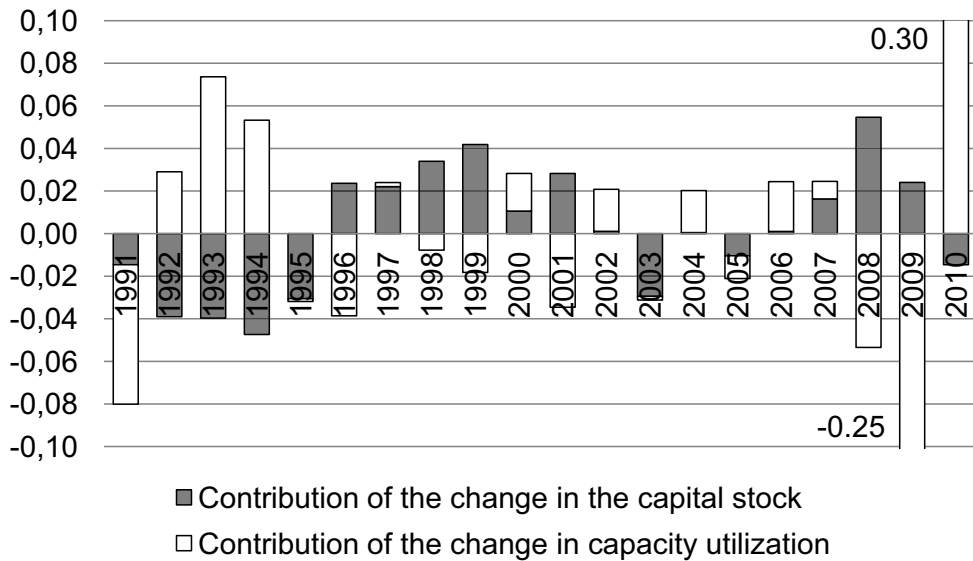
[Figure 9: Contributions of changes in the capital stock and capacity utilization to changes in active capital (cont.)]

Machinery
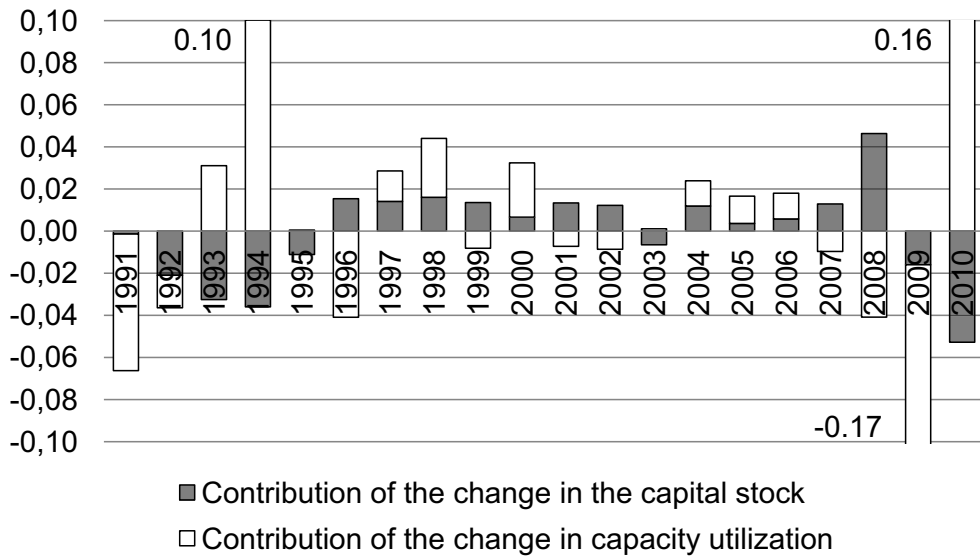
[Figure 10: Contributions of changes in the capital stock and capacity utilization to changes in active capital (cont.)]

Motor vehicles
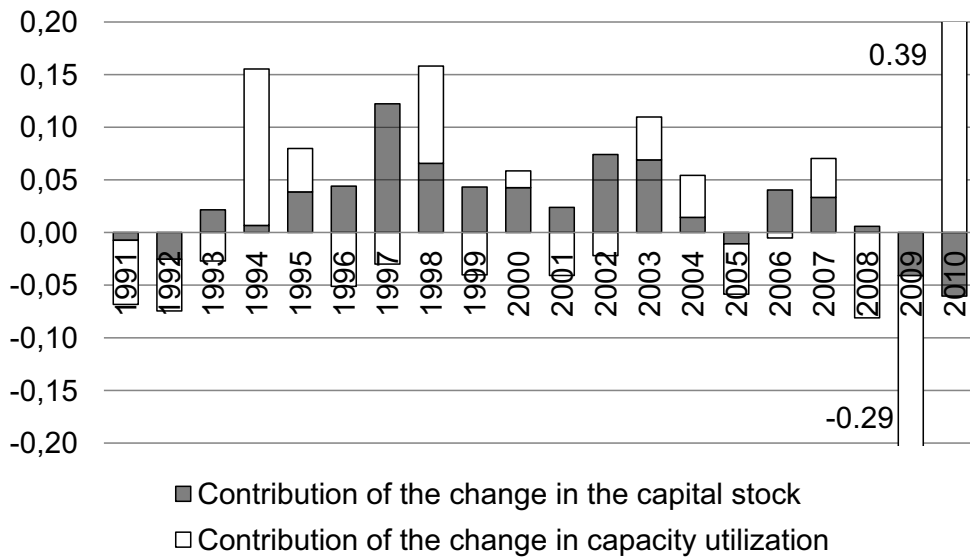
[Figure 11: Contributions of changes in the capital stock and capacity utilization to changes in active capital (cont.)]

variables		obs.	means	std. dev.	min.	max.
log of output	y	200000	0.659	1.080	-4.859	5.635
log of capital stock	k	200000	1.114	0.751	0.120	5.664
log of active capital	k^a	200000	0.747	1.071	-4.776	5.664
capacity utilization (%)	CU	200000	0.779	0.283	0.007	1.000
log of capacity utilization	cu	200000	-0.367	0.575	-4.896	0.000
sigma	σ	200000	0.202	0.059	0.100	0.300

[Table 1: Descriptive statistics for the simulated firm-level data]

variables		obs.	means	std. dev.	min.	max.
log of output	y	400	1.223	0.179	0.871	1.685
log of capital stock	k	400	1.460	0.207	1.053	2.029
log of active capital	k^a	400	1.302	0.178	0.952	1.755
capacity utilization (%)	CU	400	0.923	0.030	0.843	0.981
log of capacity utilization	cu	400	-0.080	0.033	-0.171	-0.019
sigma	σ	400	0.202	0.059	0.100	0.300

[Table 2: Descriptive statistics for the simulated sector-aggregated data]

Estimation method: fixed-effects (FE) with instrumental variables (IV)

Explanatory variables	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$
$\Delta y_{i,t}$	0.244 ***	0.693 ***	0.948 ***
	(0.003)	(0.003)	(0.001)
$(\Delta y_{i,t})^2$	0.377 ***	-0.380 ***	-0.009 ***
	(0.002)	(0.002)	(0.001)
$y_{i,t-1}$	0.010 ***	-0.008 ***	0.002 ***
	(0.000)	(0.000)	(0.000)
$k^a_{i,t-1} - y_{i,t-1}$	-0.535 ***	0.222 ***	-0.133 ***
	(0.004)	(0.003)	(0.003)
$cu_{i,t-1}$	0.002 ***	-0.011 ***	-0.006 ***
	(0.000)	(0.000)	(0.000)
$\sigma_{i,t-1}$	-0.002	0.000	0.003 ***
	(0.003)	(0.003)	(0.001)
$\sigma_{i,t-1} \cdot \Delta y_{i,t}$	-0.382 ***	0.508 ***	0.099 ***
	(0.013)	(0.014)	(0.004)
Dummy variables			
year	X	X	X
sector (2-digit)	X	X	X
Instrumental variable method			
instrumented	$k^a_{i,t-1} - y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1} - y_{i,t-1}$, $cu_{i,t-1}$
instruments	$k_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}$, $cu_{i,t-2}$
Specification tests			
underidentification test	0.000	0.000	0.000

[Table 3: Regression analysis using the simulated firm-level data]

Estimation method:	equation for $\Delta k_{i,t}$		equation for $\Delta cu_{i,t}$		equation for $\Delta k^a_{i,t}$	
	dif-GMM	dif-GMM	dif-GMM	dif-GMM	dif-GMM	dif-GMM
Explanatory variables						
$\Delta y_{i,t}$	0.206 ***	0.206 ***	0.747 ***	0.747 ***	0.961 ***	**0.961 *****
	(0.003)	(0.004)	(0.004)	(0.004)	(0.001)	**(0.001)**
$(\Delta y_{i,t})^2$	0.269 ***	0.272 ***	-0.352 ***	-0.351 ***	-0.009 ***	**-0.011 *****
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	**(0.001)**
$y_{i,t-1}$	-0.300 ***	-0.265 ***	0.046 ***	0.051 ***	0.022 ***	**0.017 *****
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	**(0.001)**
$k^a_{i,t-1}-y_{i,t-1}$	1.340 ***	1.702 ***	-0.780 ***	-0.774 ***	-0.454 ***	**-0.370 *****
	(0.013)	(0.015)	(0.010)	(0.010)	(0.007)	**(0.012)**
$cu_{i,t-1}$	0.476 ***	0.445 ***	-0.103 ***	-0.109 ***	-0.027 ***	**-0.018 *****
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	**(0.002)**
$\sigma_{i,t-1}$	0.087 ***	0.091 ***	-0.058 ***	-0.058 ***	0.009 ***	**0.010 *****
	(0.006)	(0.007)	(0.007)	(0.007)	(0.002)	**(0.002)**
$\sigma_{i,t-1}\,\Delta y_{i,t}$	-0.196 ***	-0.189 ***	0.396 ***	0.396 ***	0.069 ***	**0.075 *****
	(0.014)	(0.015)	(0.015)	(0.015)	(0.004)	**(0.005)**
Dummy variables						
year	X	X	X	X	X	X
Instrumental variable method						
instrumented	$k^a_{i,t-1}-y_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$, $cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$, $cu_{i,t-1}$
instruments	$k_{i,t-2}$	$k_{i,t-3}$	$cu_{i,t-2}$	$cu_{i,t-3}$	$k_{i,t-2}$, $cu_{i,t-2}$	$k_{i,t-3}$, $cu_{i,t-3}$
Specification tests						
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.000	0.000	0.000	0.000	0.000	0.000
AR(3)	0.000	0.000	0.000	0.000	0.000	0.120
Sargan test	0.000	0.001	0.000	0.000	0.000	0.999

[Table 4: Regression analysis using the simulated firm-level data with dif-GMM]

Estimation method: fixed-effects (FE) with instrumental variables (IV)

Explanatory variables	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$
$\Delta y_{i,t}$	0.568 ***	0.115	0.899 ***
	(0.188)	(0.077)	(0.049)
$(\Delta y_{i,t})^2$	0.641	-0.896	0.332
	(1.275)	(0.548)	(0.342)
$y_{i,t-1}$	-0.070 *	0.004	-0.014 **
	(0.040)	(0.007)	(0.007)
$k^a_{i,t-1} - y_{i,t-1}$	-2.888 **	-0.195 **	-0.947 ***
	(1.406)	(0.080)	(0.207)
$cu_{i,t-1}$	0.049	-0.159 ***	-0.058 ***
	(0.078)	(0.019)	(0.014)
$\sigma_{i,t-1}$	0.084 *	-0.099 ***	-0.021 **
	(0.049)	(0.009)	(0.009)
$\sigma_{i,t-1} \cdot \Delta y_{i,t}$	-1.036 *	0.947 ***	-0.081
	(0.603)	(0.267)	(0.165)
Dummy variables			
year	X	X	X
sector (2-digit)	X	X	X
Instrumental variable method			
instrumented	$k^a_{i,t-1} - y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1} - y_{i,t-1}$, $cu_{i,t-1}$
instruments	$k_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}$, $cu_{i,t-2}$
Specification tests			
underidentification test	0.015	0.000	0.000

[Table 5: Regression analysis using the simulated sector-aggregated data]

[Table 6: Regression analysis using the simulated sector-aggregated data with dif-GMM and sys-GMM]

Estimation method:	equation for $\Delta k_{i,t}$		equation for $\Delta cu_{i,t}$		equation for $\Delta k^a_{i,t}$	
	dif-GMM	sys-GMM	dif-GMM	sys-GMM	dif-GMM	sys-GMM
Explanatory variables						
$\Delta y_{i,t}$	**0.700 *****	**0.740 *****	**0.139**	**0.106**	**0.939 *****	**0.926 *****
	(0.110)	(0.101)	(0.091)	(0.088)	(0.041)	(0.041)
$(\Delta y_{i,t})^2$	**1.430 ***	**1.177 ***	**-0.597**	**-0.664**	**0.778 *****	**0.732 ****
	(0.741)	(0.706)	(0.620)	(0.618)	(0.282)	(0.288)
$y_{i,t-1}$	**-0.234 *****	**-0.021 ****	**0.079 *****	**0.007**	**-0.012**	**-0.001**
	(0.034)	(0.009)	(0.028)	(0.009)	(0.009)	(0.004)
$k^a_{i,t-1}-y_{i,t-1}$	**-0.969 *****	**-0.516 *****	**-0.574 *****	**-0.338 *****	**-0.743 *****	**-0.411 *****
	(0.352)	(0.182)	(0.113)	(0.092)	(0.087)	(0.061)
$cu_{i,t-1}$	**0.640 *****	**0.212 *****	**-0.408 *****	**-0.187 *****	**0.030**	**-0.017**
	(0.051)	(0.023)	(0.107)	(0.037)	(0.038)	(0.014)
$\sigma_{i,t-1}$	**0.103 *****	**0.152 *****	**-0.091 *****	**-0.098 *****	**0.006**	**0.001**
	(0.018)	(0.013)	(0.011)	(0.011)	(0.005)	(0.005)
$\sigma_{i,t-1}\Delta y_{i,t}$	**-1.308 *****	**-1.330 *****	**0.898 *****	**0.982 *****	**-0.161**	**-0.143**
	(0.388)	(0.352)	(0.320)	(0.303)	(0.145)	(0.142)
Dummy variables						
year	X	X	X	X	X	X
sector (2-digit)		X		X		X
Instrumental variable method						
instrumented	$k^a_{i,t-1}-y_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$
instruments	$k_{i,t-2}$	$k_{i,t-2}$	$cu_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}, cu_{i,t-2}$	$k_{i,t-2}, cu_{i,t-2}$
Specification tests						
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.192	0.015	0.924	0.916	0.072	0.204
Sargan test	0.273	0.022	0.265	0.182	0.053	0.000
Dif-in-Sargan test		0.003		0.102		0.000

variables		obs.	means	std. dev.	min.	max.
output (million Swedish crowns)	Y	319	27631	34424	1625	281092
capital stock (million Swedish crowns)	K	319	28422	21696	2568	84529
active capital (million Swedish crowns)	K^a	319	25174	19729	1842	80940
capacity utilization (%)	CU	319	0.873	0.054	0.634	1.000
log of output	y	319	9.775	0.968	7.393	12.546
log of capital stock	k	319	9.928	0.864	7.851	11.345
log of active capital	k^a	319	9.798	0.873	7.519	11.301
Log of capacity utilization	cu	319	-0.138	0.064	-0.456	0.000

[Table 7: Descriptive statistics for the real data]

Estimation method: fixed-effects (FE) with instrumental variables (IV)

	σ (uncertainty) is constructed with *production volume*			σ (uncertainty) is constructed with *order inflows*			σ (uncertainty) is constructed with *selling prices*		
Explanatory variables	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$
$\Delta y_{i,t}$	0.062 **	0.125 ***	0.162 ***	-0.043	0.132 ***	0.065	0.002	0.099 ***	0.090 **
	(0.029)	(0.044)	(0.054)	(0.033)	(0.050)	(0.060)	(0.025)	(0.036)	(0.044)
$(\Delta y_{i,t})^2$	0.087 ***	-0.223 ***	-0.107 **	0.101 ***	-0.226 ***	-0.100 **	0.096 ***	-0.293 ***	-0.171 ***
	(0.027)	(0.043)	(0.052)	(0.027)	(0.042)	(0.050)	(0.030)	(0.044)	(0.053)
$y_{i,t-1}$	-0.044 ***	0.026	-0.026	-0.053 ***	0.022	-0.034 *	-0.049 ***	0.023	-0.035 *
	(0.011)	(0.019)	(0.020)	(0.011)	(0.019)	(0.020)	(0.011)	(0.018)	(0.020)
$k^a_{i,t-1}-y_{i,t-1}$	-0.043 ***	0.019	-0.029	-0.050 ***	0.015	-0.036 *	-0.046 ***	0.018	-0.036 *
	(0.012)	(0.020)	(0.021)	(0.012)	(0.020)	(0.021)	(0.012)	(0.019)	(0.021)
$cu_{i,t-1}$	0.144 ***	-0.530 ***	-0.257 *	0.184 ***	-0.527 ***	-0.239 **	0.160 ***	-0.574 ***	-0.295 ***
	(0.034)	(0.109)	(0.127)	(0.034)	(0.100)	(0.117)	(0.033)	(0.095)	(0.111)
$\sigma_{i,t-1}$	-0.006	-0.036	-0.025	0.035	-0.096 **	-0.052	-0.018	-0.044	-0.062
	(0.022)	(0.034)	(0.042)	(0.027)	(0.040)	(0.050)	(0.026)	(0.039)	(0.048)
$\sigma_{i,t-1}\cdot\Delta y_{i,t}$	-0.226 **	0.244 *	0.097	0.141	0.208	0.411 **	0.001	0.486 ***	0.506 ***
	(0.093)	(0.144)	(0.177)	(0.100)	(0.150)	(0.181)	(0.108)	(0.152)	(0.186)
Dummy variables									
year	X	X	X	X	X	X	X	X	X
sector (2-digit)	X	X	X	X	X	X	X	X	X
Instrumental variable method									
instrumented	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$
instruments	$k_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}, cu_{i,t-2}$	$k_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}, cu_{i,t-2}$	$k_{i,t-2}$	$cu_{i,t-2}$	$k_{i,t-2}, cu_{i,t-2}$
Marginal effects of $\Delta y_{i,t}$ at different values of σ									
σ (10% point)	0.044 **	0.138 ***	0.166 ***	-0.022	0.144 ***	0.106 **	0.006	0.118 ***	0.117 ***
	(0.021)	(0.032)	(0.039)	(0.023)	(0.035)	(0.042)	(0.019)	(0.028)	(0.034)
σ (50% point)	0.024	0.160 ***	0.175 ***	-0.010	0.162 ***	0.141 ***	0.006	0.143 ***	0.142 ***
	(0.016)	(0.023)	(0.028)	(0.017)	(0.025)	(0.031)	(0.016)	(0.023)	(0.028)
σ (90% point)	-0.004	0.190 ***	0.187 ***	0.010	0.192 ***	0.199 ***	0.007	0.212 ***	0.215 ***
	(0.014)	(0.020)	(0.026)	(0.014)	(0.020)	(0.025)	(0.015)	(0.021)	(0.026)

[Table 8: Regression analysis using the real data]

Estimation method: dif-GMM

Explanatory variables	σ (uncertainty) is constructed with *production volume*			σ (uncertainty) is constructed with *order inflows*			σ (uncertainty) is constructed with *selling prices*		
	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$	equation for $\Delta k_{i,t}$	equation for $\Delta cu_{i,t}$	equation for $\Delta k^a_{i,t}$
$\Delta y_{i,t}$	0.008	**0.101** **	0.079 *	-0.020	**0.146** ***	0.078 *	0.048 ***	**0.030**	0.064 *
	(0.021)	**(0.044)**	(0.042)	(0.021)	**(0.043)**	(0.041)	(0.019)	**(0.034)**	(0.037)
$(\Delta y_{i,t})^2$	0.038	**-0.151** ***	-0.087 *	0.044 *	**-0.539** ***	-0.090 *	0.042 *	**-0.206** ***	-0.158 ***
	(0.024)	**(0.051)**	(0.048)	(0.023)	**(0.077)**	(0.047)	(0.024)	**(0.046)**	(0.048)
$y_{i,t-1}$	-0.375 ***	**-0.558** ***	-0.404 ***	-0.385 ***	**-0.162** ***	-0.420 ***	-0.373 ***	**-0.364** ***	-0.402 ***
	(0.052)	**(0.076)**	(0.093)	(0.051)	**(0.051)**	(0.093)	(0.052)	**(0.095)**	(0.092)
$k^a_{i,t-1}-y_{i,t-1}$	-0.420 ***	**-0.643** ***	-0.505 ***	-0.430 ***	**-0.630** ***	-0.524 ***	-0.420 ***	**-0.437** ***	-0.500 ***
	(0.056)	**(0.080)**	(0.102)	(0.055)	**(0.081)**	(0.102)	(0.056)	**(0.098)**	(0.101)
$cu_{i,t-1}$	0.495 ***	**-0.171**	-0.394 ***	0.509 ***	**-0.225**	-0.436 ***	0.477 ***	**-0.519** ***	-0.456 ***
	(0.060)	**(0.148)**	(0.123)	(0.059)	**(0.149)**	(0.124)	(0.060)	**(0.160)**	(0.120)
$\sigma_{i,t-1}$	0.001	**0.055** *	0.057 **	0.019	**-0.057** *	-0.056 **	-0.023	**-0.042**	-0.073 **
	(0.013)	**(0.029)**	(0.027)	(0.014)	**(0.030)**	(0.028)	(0.017)	**(0.031)**	(0.033)
$\sigma_{i,t-1} \cdot \Delta y_{i,t}$	0.062	**0.261** **	0.290 **	0.135 **	**0.062**	0.218 **	-0.126 *	**0.638** ***	0.446 ***
	(0.064)	**(0.133)**	(0.127)	(0.053)	**(0.111)**	(0.106)	(0.075)	**(0.142)**	(0.154)
Dummy variables									
year	X	X	X	X	X	X	X	X	X
sector (2-digit)	X	X	X	X	X	X	X	X	X
Instrumental variable method									
instrumented	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}$	$cu_{i,t-1}$	$k^a_{i,t-1}-y_{i,t-1}, cu_{i,t-1}$
instruments	$k_{i,t-3}$	$cu_{i,t-2}$	$k_{i,t-3}, cu_{i,t-3}$	$k_{i,t-3}$	$cu_{i,t-2}$	$k_{i,t-3}, cu_{i,t-3}$	$k_{i,t-3}$	$cu_{i,t-3}$	$k_{i,t-3}, cu_{i,t-3}$
Specification tests									
AR(1)	0.208	0.000	0.062	0.136	0.000	0.085	0.314	0.000	0.127
AR(2)	0.703	0.184	0.313	0.493	0.082	0.436	0.395	0.745	0.105
AR(3)	0.002		0.183	0.001		0.141	0.001	0.264	0.171
Sargan test	0.001	0.167	0.000	0.000	0.145	0.000	0.001	0.228	0.000

[Table 9: Regression analysis using the real data with dif-GMM]

Paper 3

Are Larger Firms More Productive because of Scale Economies? – Evidence from Swedish Register Data[*]

Måns Söderbom and Yoshihiro Sato
University of Gothenburg

Abstract

This study investigates the factors driving higher labor productivity for large firms, using Swedish register-based microdata for the mining and manufacturing industries covering more than 28,000 firms during 1997-2006. We estimate translog production functions using dynamic panel approaches and the approach proposed by Ackerberg, Caves, and Frazer. The results show that micro and small firms operate under (locally) increasing returns to scale while medium and large firms face decreasing returns to scale. Scale elasticity decreases from 1.15 to 0.97, suggesting that scale effects are not the answer to our question. Further investigation shows that production technology is approximated by a non-homothetic function and that larger firms operate with more capital-intensive technology while the factor price ratio is constant, which drives the productivity difference in favor of larger firms.

1 Introduction

In a world where economies are becoming increasingly globalized, it is a crucial question for developed countries how to survive in the competition with catching-up economies. Issues such as the importance of productivity growth, new innovative firms, and entrepreneurship have been widely discussed in the policy debate in recent decades. In the Swedish context, the Globalisation Council, established by the government, stresses the importance of small firms and entrepreneurship for the future prosperity of the economy in a report on economic policy strategies (Globalisation Council, 2009a).

[*]I am particularly grateful to the comments and the STATA codes provided by Dario Pozzoli at University of Aarhus. I also appreciate the comments by Dan Johansson, Sven-Olov Daunfeldt, and other participants in a seminar at the Ratio Institute in Stockholm.

Previous studies, however, report that small firms are less efficient and productive than larger firms (e.g., Brouwer et al., 2005; Castany et al., 2005; Taymaz, 2005). As we will show, Swedish firm-level data from the mining and manufacturing industries also indicate that larger firms on average are more productive than smaller ones.

The purpose of this study is to explore the seemingly contradictory fact that, while it is commonly argued that new and small firms play an important role in economic growth, it is the larger firms that exhibit the highest productivity. More specifically, we investigate the relative importance of possible factors driving the productivity difference between firms of different sizes by decomposing labor productivity into contributions of economies of scale, variation in capital intensity, and individual firm effects.

Understanding the factors behind the productivity difference between firms of different size is important in order to answer relevant policy questions, e.g., whether firms of certain sizes have larger growth potentials than firms of other sizes in terms of productivity and efficiency. In an economy with a Cobb-Douglas production technology with constant returns to scale, there are no benefits from reallocating resources from small to large firms or vice versa. However, when production technology is no longer characterized by constant returns to scale, it is possible to obtain efficiency gains through more efficient allocation of resources. One example of such production technology is technology characterized by a non-homothetic production function.

Our study is unique in that we use register-based panel data for the Swedish mining and manufacturing industries. Our dataset covers 10 years from 1997 to 2006 and follows more than 28,000 firms with 177,000 observations. The dataset is extracted from the Structural Business Statistics. It therefore covers almost all firms active in these industries in Sweden and is, unlike previous studies, dominated by micro firms.

The results show that scale elasticity is decreasing with firm size – while micro and small firms operate under (locally) increasing returns to scale, medium and large firms face (locally) decreasing returns to scale. This indicates that the answer to the question of why larger firms are more productive is not economies of scale. On the contrary, decomposition of labor productivity reveals that production technology is expressed by a non-homothetic function and that this causes capital intensity to increase with firm size whereas the factor price ratio is homogeneous. An increase in capital intensity raises labor productivity more than a decrease in economies of scale lowers it. This explains the positive

association between firm size and labor productivity.

As estimation methods, we apply different approaches: a fixed-effect (within-effect) model, a dynamic panel model with the generalized method of moments (GMM), and the approach proposed by Ackerberg et al.(2006). The comparison shows the features and, if anything, the shortcomings of these approaches. We also address the issue of market power, which may bias the estimated scale elasticity. To tackle this problem, we also estimate the empirical model using the approach proposed by Klette and Griliches (1996).

The paper is organized as follows. Section 2 investigates discussions on the relationship between productivity and firm size, and on the emerging importance of entrepreneurship in economic growth. Section 3 describes the register data used in the empirical analysis. Section 4 estimates translog production functions and derives scale elasticity. Section 5 discusses potential biases generated by imperfect product markets and suggests and implements a partial solution to the problem. Section 6 checks the robustness of the result at the aggregate level by estimating the production function at the sectoral level. Section 7 discusses decomposition of labor productivity to explain the productivity difference across different firm sizes. Section 8 concludes the paper.

2 Background

2.1 Recognition of the importance of entrepreneurship and small and medium-sized firms

Until the early 1970s, it was widely accepted that economic progress could be achieved through mass and specialized production of standardized products in capital-intensive industries. Large-scale enterprises were therefore considered an economic engine, because they were supposed to be best suited to realize gains from economies of scale.

However, the oil crises in the 1970s became a turning point for this growth story. The long-standing trend toward centralization in the business organization ceased and reversed, and the share of employment in small and medium-sized enterprises and establishments began to increase in many developed countries. Loveman and Sengenberger (1991) provide evidence using data from the six largest OECD countries (United States, France, Germany, United Kingdom, Italy, and Japan), and explore various explanations for why the size advantages seem to diminish. They argue that small and medium-sized firms often enjoy

lower labor costs and flexibility in specialization, while large firms are often subject to inflexibility in adapting to changing market conditions. As a result of the reduced importance of economies of scale, large firms are likely to downsize or outsource part of their internal activities to concentrate on their core competence (Acs and Audretsch, 1989). Carlsson (1992) argues that changes in the economic environment (e.g., the intensified global competition and an increase in the degree of uncertainty) and technological progress (e.g., flexible automation) have resulted in a structural shift from large to small firms.[1]

Alongside the increasing importance of small and medium-sized firms, the role of new firms in economic growth and job creation has also been recognized (Davis and Haltiwanger, 1992). Geroski (1995) argues that entrepreneurial activities play an important role in the evolution of market structure and market performance by, for example, promoting new product innovation, discovering new markets, and replacing inefficient incumbents. He also points out that entry of new firms leads to erosion of high profits among incumbents and lower prices for consumers. This is the process called "creative destruction" (Schumpeter, 1942) and is expected to enhance economic growth.

The Swedish economy has traditionally been dominated by many large and relatively few small firms compared to other European countries. Davis and Henrekson (1997a) and Henrekson and Johansson (1999) report that, while large firms had already begun to lose importance in other industrial economies since the 1970s, their dominance remained or even increased in Sweden through the first half of the 1980s. In an attempt to explain this observation, Davis and Henrekson (1997b), Henrekson and Johansson (1999, 2009), and Henrekson (2005) argue that the institutional conditions have been unfavorable for small firms, and therefore too few small firms have managed to grow out of the smallest size classes.[2]

Thus, the focus on promotion of small/medium-sized firms and entrepreneurial activities by the Swedish governmental Globalization Council can be interpreted as an ambition to reverse this trend and stimulate the segments of smaller and newer firms by identifying key institutional settings. A similar political ambition is observed in other countries. For example, the European Union adopted the European Charter for Small Enterprises in June 2000 as part

---

[1] Audretsch (2005) explaines the shift of the view on entreprenuership.

[2] These conditions include the tax system, regulated credit market conditions, employment security laws, low wage dispersion due to wage-setting institutions, and the public sector monopolizing the production of key services.

of the Lisbon strategy, which promotes the idea of the European Union will becoming the most competitive, dynamic, and knowledge-intensive economy by 2010 (European Council, 2000).

2.2 Firm size and productivity

Most previous studies, however, cast doubt on the hypothesis that smaller firms are more efficient and productive. Evidence on the size-efficiency relationship indicates that there is a positive association between firm size and technical efficiency, and there are also substantial and persistent productivity differences between small/medium-sized firms and large firms (Brouwer et al., 2005; Castany et al., 2005; Taymaz, 2005).[3] Furthermore, the exit rate of small and young firms is generally high (Dunne et al., 1988), and Italy, Spain, and Portugal – countries that have recorded poor economic performance in recent years – are known for domination of relatively many small firms and few large firms (Henrekson and Johansson, 1999; Stenkula, 2006). They were also the three countries in Western Europe with the highest number of newly started firms per capita in 2006.[4]

There are some explanations for the findings that small and medium-sized firms are often less efficient. Cohen and Levin (1989) and Alvarez and Crespi (2003) point to, for example, their inability to take advantage of scale economies, their difficulties they face in getting access to credit for investment due to the riskiness of R&D projects and capital market imperfections, and their lack of resources in terms of qualified human capital and knowledge networks. Jovanovic (1982) instead focuses on the selection process of the industry dynamics. He argues that larger firms are more efficient because only efficient firms survive to grow large, while inefficient firms fail to grow and hence remain small.

Another explanation could be that their is a link between firm size and innovation. While innovative activities are supposed to be an important way for new firms to successfully compete with incumbent firms (Acs and Audretsch, 1989; Geroski, 1999), some researchers report that new and therefore often small firms engage in innovative activities to a less extent than do larger firms. For instance, Cohen (1995) reviews empirical studies in the field and concludes that there is a positive relationship between market concentration and R&D. Cohen and Klepper (1996) find that net innovative output is increasing with firm size.

---

[3] For a comprehensive review, see Yang and Chen (2009).
[4] According to the data from Eurostat presented in Globalisation Council (2009b).

As an explanation, they point to the advantages of large firms in innovation as large firms are able to reduce duplicative effort and undertake projects that would not be profitable for small firms. Using data from Swedish manufacturing firms, Nyström (2005) finds a positive association between R&D activities and the age and size of firms. Her finding supports the argument that innovation constitutes a barrier to entry (Bain, 1956 and Yip, 1982).

However, there are also empirical results that imply the opposite. Baumol (2004) argues that many new innovations in the US have been developed in small and medium firms. Geroski (1995) shows that small firms and entering firms make a substantial contribution to the generation and diffusion of innovation.[5]

3   Data and descriptive statistics

3.1   Dataset

Our empirical study is based on firm-level panel data from the Structural Business Statistics from Statistics Sweden. The original database contains detailed information on the income statements, balance sheets, and physical investment of all firms active in Sweden, including private and public firms but not financial firms. Most of the data are obtained from registers at the Swedish national tax agency. All firms are classified according to NACE (Classification of Economic Activities in the European Community) with 5-digit classes.[6] The dataset contains data for manufacturing and mining firms from 1997 to 2006.[7]

The number of firms in the extracted database is originally 101,515 with 556,930 observations. However, it turns out that more than half of the firms have no employees (e.g., self-employed). We drop all firms without employees in some or all years in the period 1997-2006. As a result, our unbalanced panel dataset contains 28,313 firms with a total of 177,649 observations (the average number of observations per firm is thus 6.27).

The original dataset contains a book value of the capital stock for each firm.

---

[5] A comprehensive review of different arguments regarding the relationship between R&D activities and firm size and age is found in Nyström (2005).

[6] However, some industries are only classified at the 3-digit or the 4-digit level at most. The number of subsectors at the 5-digit level is 294, while the numbers at the 4-digit and the 3-digit levels are 253 and 113, respectively.

[7] Although the majority of the infomation in the database is register based instead of survey based, there are some missing firms. As a proportion of the whole population in manufacturing and mining industries, the missing firms amount to 12.6 percent. However, if we weight them with the number of employees and sales, the proportion decreases to 1.7 and 1.3 percent, respectively.

However, a book value may divert from the real value of the capital stock for tax-related or other reasons. We therefore rely on investment data and apply the perpetual inventory method to create a series of the capital stock, with a start value set to the book value in the first year the firm appears in the dataset. Although the dataset also contains data on land as a component of physical capital, we define the capital stock as the sum of only buildings and machinery.

The number of employees in the dataset is the annual average converted into full-time equivalents. Intermediate inputs and the capital stock are deflated with the respective deflators at the 2-digit sector level, derived from the Swedish national account. For value added, there is a detailed price index available from the price division of Statistics Sweden. This price index is only defined at the 2-digit level for some industries, while it is defined at the 3-digit or 4-digit level for other industries. The number of subsectors for which the price index is defined is 141.

The database used, the Structural Business Statistics in Sweden, has previously been used in other empirical studies, though their sample periods and industries/sectors vary. For research on the relationship between wage and ownership, see Heyman et al.(2007). For research on the relationship between productivity/innovation and other key factors (e.g., FDI, spillover, ownership, exporters), see for example Karpaty (2004), Karpaty and Lundberg (2004), Ekholm and Hakkala (2005), Nyström (2005), Poldahl (2006), Andersson and Karpaty (2007), Lööf (2008), Andersson et al.(2008), Oh et al.(2009a), Eliasson et al.(2009), Andersson and Lööf (2009a, 2009b, 2010), and Bjuggren et al.(2010).

3.2   Descriptive statistics by firm size

The European Commission has several definitions for classification of firms into different size groups. One is based on number of employees.[8]   According to this definition, a firm qualifies as micro, small, and medium-sized if the number of employees is less than 10, between 10 and 49, and between 50 and 249, respectively. Otherwise, it is classified as a large firm.

Table 1 presents some key variables in our dataset that will be used in the empirical analysis. As the first two rows show, a majority of firms are micro firms. However, when we turn to the shares of value added, employment, and

---

[8] The other definitions are based on firms' turnover and asset (balance sheet total). http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/index_en.htm

the capital stock, a majority belongs to large firms while only a small portion belongs to micro firms.

[Table 1: Summary statistics by firm size]

The table also reports logs of value added per employee (i.e., labor productivity) and the capital stock per employee (i.e., capital intensity) by size group. We report the conditional means based on OLS regressions including dummy variables for sector, year, and firm size[9] to separate year-specific and sector-specific effects. They have clear trends in favor of large firms, which indicates that larger firms have higher labor productivity and capital intensity.

Table 2 presents the size distribution of firms at the 2-digit sectoral level. There are four industries (pulp and paper; coke, petroleum and other fuels; chemicals; and motor vehicles) where medium-sized and large firms constitute more than 20%. In other industries, micro and small firms dominate.[10]

[Table 2: Distribution of firm size in different sectors]

4   Conceptual framework

4.1   Decomposition of productivity

We have observed that labor productivity (value added per worker) is higher for larger firms than for smaller firms. Our question is whether this is a result of increasing returns to scale or of other factors such as higher capital intensity (capital stock per worker) in larger firms. To understand the relationship between capital intensity, increasing returns to scale, and productivity, we first consider a Cobb-Douglas production function

$$y_{i,t} = \alpha_l l_{i,t} + \alpha_k k_{i,t} + (\lambda_t + \eta_i + v_{i,t}),\tag{1}$$

where α_l and α_k are output elasticities with respect to labor and capital, respectively, λ_t is a year-specific effect (e.g., a common productivity or demand shock),

---

[9] That is, $\hat{a} \cdot D_{size\_i} + \hat{\mathbf{b}} \cdot \bar{\mathbf{x}}$, where D_{size_i} indicates the dummy for the i-th size category, $\bar{\mathbf{x}}$ is the vector of mean values of the sector and time dummies, and $\hat{a}$ and $\hat{\mathbf{b}}$ are estimated coefficients.

[10] The firm size is defined as the mean firm size in the sample period, since some firms move across different size groups. For a similar reason, we also define the industry classification of a firm as the one it belongs to the first time it appears in the dataset.

η_i is an unobserved time-invariant firm-specific effect (e.g., some characteristics of firms such as management and technology), and $v_{i,t}$ is a residual that is not necessarily white noise. Then, the log of labor productivity is expressed by

$$y_{i,t} - l_{i,t} = (\alpha_l + \alpha_k - 1) \, l_{i,t} + \alpha_k \, (k_{i,t} - l_{i,t}) + \lambda_t + \eta_i + v_{i,t}, \qquad (2)$$

implying that labor productivity can be decomposed into three components: (i) deviation from constant returns to scale $(\alpha_l + \alpha_k - 1)$, (ii) capital intensity $(k_{i,t} - l_{i,t})$, and (iii) individual firm effects η_i.

The decomposition is complicated in the case of a translog production function. Assuming the functional form

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \frac{1}{2}\beta_{ll} l_{i,t}^2 + \beta_{lk} l_{i,t} k_{i,t} + \frac{1}{2}\beta_{kk} k_{i,t}^2 + \lambda_t + \eta_i + v_{i,t}, \qquad (3)$$

the log of labor productivity is decomposed as

$$
\begin{aligned}
y_{i,t} - l_{i,t} \;=\; & \left(\frac{\partial y_{i,t}}{\partial l_{i,t}} + \frac{\partial y_{i,t}}{\partial k_{i,t}} - 1\right) l_{i,t} \\
& + \beta_k \, (k_{i,t} - l_{i,t}) + \frac{1}{2}\beta_{kk} \, (k_{i,t} - l_{i,t})^2 \\
& - \frac{1}{2}\left[(\beta_{ll} + \beta_{lk}) + (\beta_{lk} + \beta_{kk})\right] l_{i,t}^2 \\
& + \lambda_t + \eta_i + v_{i,t},
\end{aligned}
\qquad (4)
$$

where output elasticities are expressed as

$$\frac{\partial y_{i,t}}{\partial l_{i,t}} \;=\; \beta_l + \beta_{ll} l_{i,t} + \beta_{lk} k_{i,t} \qquad (5)$$

$$\frac{\partial y_{i,t}}{\partial k_{i,t}} \;=\; \beta_k + \beta_{kk} k_{i,t} + \beta_{lk} l_{i,t}. \qquad (6)$$

Thus, the decomposition has additional terms. The first term in Eq.(4) is interpreted as a component related to (i) deviations from constant returns to scale. The second and third terms are due to (ii) variation in capital intensity. The fourth term is something that does not exist in the decomposition of a Cobb-Douglas function. Since we have $\beta_{ll} + \beta_{lk} = 0$ and $\beta_{lk} + \beta_{kk} = 0$ under the assumption of homotheticity, we can interpret this term as a component related to (iii) deviation from homotheticity.[11] Lastly, η_i is (iv) individual firm effects.

---

[11] Note that this term can be zero although homotheticity is violated. This is when the sum of $\beta_{ll} + \beta_{lk}$ and $\beta_{lk} + \beta_{kk}$ is zero.

The implications of these components are described in Figures 1–4. Each figure exhibits isoquants for output $\overline{Y}$ and output $2\overline{Y}$. Figure 1 shows a production function that satisfies homotheticity and constant returns to scale. When the first component in Eq.(4) is non-zero, constant returns to scale is no longer valid. The isoquants of a production function with decreasing returns to scale are exhibited in Figure 2.

[Figure 1-4: Isoquants under different conditions]

The implications of (ii) capital intensity and its relationship to the factor price ratio is clear under a homothetic production function (i.e., $\beta_{ll} + \beta_{lk} = 0$ and $\beta_{lk} + \beta_{kk} = 0$). Constant capital intensity suggests a homogenous factor price ratio, whereas variable capital intensity suggests a heterogenous factor price ratio.[12] Figure 3 illustrates how a heterogenous factor price ratio affects capital intensity under a homothetic production function. Note that returns to scale are not necessarily constant.

In contrast, however, when the production function is non-homothetic so that $\beta_{ll} + \beta_{lk} \neq 0$ and/or $\beta_{lk} + \beta_{kk} \neq 0$, the relationship between capital intensity and the factor price ratio is not straightforward. Variable capital intensity is not solely driven by a heterogenous factor price ratio, but also by non-homotheticity. In other words, capital intensity can vary even when the factor price ratio is constant. Figure 4 illustrates how a non-homothetic production function affects capital intensity under a homogenous factor price ratio. When capital intensity varies under non-homotheticity, however, the effects of (ii) capital intensity and (iii) non-homotheticity on labor productivity are not separable: the effect of (ii) variable capital intensity is also canalized through (iii) non-homotheticity. To analyze the relative strengths of the effects of a non-homothetic feature of the production function and a heterogenous factor price ratio on the capital intensity, we need to investigate the factor price ratio implied by the estimated production function.

Finally, (iv) the individual firm effects are referred to as the firm-specific time-invariant component that is typically captured by fixed effects in panel models.

---

[12]See discussion in Söderbom and Teal (2004).

4.2   Scale elasticity in previous studies

Before moving to estimation of the production function, we shortly review previous studies that compute scale elasticity by estimating production functions using firm-based datasets. Many of the recent studies report scale elasticity of around 1 or slightly under 1. Some of the examples are Blundell and Bond (2000), who analyze R&D-performing US manufacturing firms using OLS and the system GMM, and Olley and Pakes (1996), who analyze US telecommunication equipment firms using OLS and their own (OP) approach.[13]

Although some of the studies referred in Footnote 13 assume a translog production function instead of a Cobb-Douglas function, few of them estimate scale elasticity at different production levels or firm sizes, in addition to scale elasticity at the mean or the median.

One exception is Badunenko et al.(2008), who analyze German manufacturing firms using a translog specification with fixed effects. The estimated scale elasticity varies only slightly from 0.995 for the smallest firm (1st percentile) to 1.020 for the largest firm (99th percentile).

Another exception is Oh et al.(2009a). Their study is especially interesting here because they investigate Swedish firms using the same database, although their study includes not only manufacturing but also services (the sample size is 5,893 firms with 37,800 observations). However, since their sample period 1992–2000 is earlier than ours, the database does not cover all firms but mostly large firms, which is obvious from the mean number of employees of 159 (the median is 44)[14] compared to our 33 (the median is 6). Output elasticities with respect to capital and labor are estimated to be 0.19 and 0.82, respectively (standard errors are 0.08 for both), and scale elasticity is thereby 1.01 (standard error is 0.02). Studying each size group, they find that output elasticity with respect

---

[13] Some other examples are: Söderbom and Teal (2004) - Ghanian manufacturing firms using system GMM; Rizov and Walsh (2009) - UK manufacturing firms using the OP approach; Hansen (2010) - German and Austrian large firms using Levisohn and Petrin's (2003) approach (LP); Vial (2006) - Indonesian manufacturing firms using LP; De Loecker (2010) - Slovenian manufacturing firms using Ackerberg, Caves, and Frazer's (2006) approach (ACF); Graham (2001) - British manufacturing firms using seemingly unrelated system equations with full information maximum likelihood (FIML); Van Leeuwen and Klomp (2006) - Dutch innovating firms using system equations with 3SLS; Bigsten et al.(2004) - sub-Saharan African manufacturing firms using the maximum likelihood method and GMM; Pavcnik (2002) - Chilean manufacturing plants using OP; Ackerberg, Caves, and Frazer (2006) - Chilean manufacturing plants using LP and ACF. All of these studies report scale elasticity of around 1 or slightly under 1, except the last study, which reports scale elasticity of 1.1–1.3.

[14] Only 1% of the firms have less than 10 employees (the authors call them micro firms), 55% have 10–50 employees (small firms), 35% have 50–300 employees (small-medium and medium firms), and 9% have more than 300 employees (large firms).

to capital is increasing with firm size while that for labor is decreasing (0.20 and 0.81 for micro firms and 0.24 and 0.79 for large firms). As a result, scale elasticity is slightly higher for larger firms.[15]

The same authors report a similar investigation using a longer panel of Korean manufacturing firms for the period 1987–2007 (Oh et al., 2009b). The estimated elasticities with respect to capital and labor are more in favor of capital compared to the Swedish case: 0.34 and 0.66. Scale elasticity is 1.00 with standard errors of 0.05. Both capital and labor elasticities are increasing with firm size, resulting in increasing scale elasticity (from 0.94 for micro firms to 1.04 for large firms). Again, the dataset is dominated by medium-sized and large firms[16].

Yet another previous study, Heshmati (2001), explores the growth patterns of Swedish micro and small firms (1 to 100 employees) for the period 1993-1998, although he does not estimate a production function. He finds a positive scale effect that improves labor productivity.

5 Estimation of the production function

5.1 Model specification and endogeneity problems

We assume a translog production function. A simpler alternative might be a Cobb-Douglas production function with or without the restriction of constant returns to scale. However, because our purpose in this study is to investigate whether and how scale elasticity varies with firm size, we choose to adapt a more flexible, translog type of production function:

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_{ll} l_{i,t}^2 + \beta_{lk} l_{i,t} k_{i,t} + \beta_{kk} k_{i,t}^2 + (\lambda_i + v_{i,t}) + e_{i,t}, \qquad (7)$$

where $(\lambda_i + v_{i,t})$ is the unobserved productivity component that is not captured by production factors. It consists of time-invariant firm-specific effects λ_i (e.g.,

---

[15] One may wonder whether the inclusion of the service sectors affects the estimated elasticities. Although they do not estimate the production function for each sector, they use sector-specific dummy variables that interact with factor inputs and time trends. Based on their model, it is shown that the service sectors (wholesale and retail; transport and communication; renting and business activities) tend to have slightly lower elasticity for capital and slightly higher elasticity for labor. At the same time, scale elasticity is almost the same as manufacturing sectors.

[16] The sample size is 7,462 firms with 60,900 observations. The size composition is as follows: 1% of the firms have less than 10 employees (micro firms), 25% have 10–50 employees (small firms), 56% have 50–300 employees (medium-sized firms), and 18% have over 300 employees (large firms).

some characteristics of firms such as management and technology) and a serially correlated residual $v_{i,t}$. The last part $e_{i,t}$ stands for a white noise, or we call it "measurement error" that is serially uncorrelated (Blundell and Bond, 2000).

5.1.1 Dynamic panel model

Three problems in estimating a production function in a panel model are widely discussed in previous studies. The following is a short review of the problems and the proposed solutions.

(1) Treatment of unobserved time-invariant firm-specific λ_i

An OLS estimator is expected to be biased because the firm-specific effect λ_i, which is included in the OLS residual, is likely to positively correlate with the regressors. One solution to control for the unobserved individual heterogeneity is a fixed-effect (within-group) model and a first-difference model.

(2) Endogeneity of some of the regressors

Although a fixed-effect model and a first-difference model separate the firm-specific effect λ_i from an idiosyncratic shock $v_{i,t}$, another problem arises when regressors are correlated with $v_{i,t}$. This is the case if $v_{i,t}$ is unobservable for researchers but observable for the firm's manager, who adjusts factor inputs once realization of $v_{i,t}$ is known. One way to solve this problem is to include lagged regressors as instrumental variables in a first-difference model,[17] which is estimated with the Generalized Method of Moment (GMM). This is proposed by Arellano and Bond (1991) and known as "difference GMM." However, lagged regressors perform poorly as instruments and cause large finite-sample biases, when the series are close to unit root, which is the case for our dataset.[18] Arellano and Bover (1995) and Blundell and Bond (1998) propose to estimate the first-difference equation together with the level equations[19] (known as "system GMM") to reduce the biases.

---

[17] This method does not work in a fixed-effect model, because all realizations of the disturbances are introduced into the error term of the transformed equation, and the error term therefore correlates with any lag of the regressors.

[18] Simple AR(1) specifications of the time series of labor, capital, and value added shows that the autoregressive coefficients are 0.724 (0.012) for labor, 0.901 (0.008) for capital, and 0.628 (0.012) for value added (standard errors in parantheses). The model used is $x_{i,t} = \rho_x x_{i,t-1} + \pi_t + \epsilon_{i,t}$, where ρ_x is the autoregressive coefficient, π_t is a time-specific effect and $\epsilon_{i,t}$ is white noise. These estimates are obtained using a system GMM model with the second and third lags as instruments. When the third and fourth lags are used as instruments, the estimates change to 0.956 (0.011), 0.891 (0.009), and 0.995 (0.015). The persistency seems to be quite large, indicating that lagged regressors are only weakly correlated with subsequent first differences.

[19] The regressors in levels are instrumented with lags of their own first-differences.

(3) Autocorrelation in the unobserved productivity shock $v_{i,t}$

Yet another concern is potential autocorrelation in the unobserved idiosyncratic productivity shock $v_{i,t}$. The moment conditions in the GMM estimation are only valid under the assumption of no serial correlation in the error term. In the context of production functions, however, a firm with a positive short-term productivity shock is likely to have a positive shock even in the following period, namely

$$v_{i,t} = \rho v_{i,t-1} + u_{i,t} \qquad (0 \leq \rho \leq 1). \tag{8}$$

A solution is to put Eq.(7) and its lagged equation into Eq.(8) to make a common factor model:

$$
\begin{aligned}
y_{i,t} =\ & \rho y_{i,t-1} + \beta_l l_{i,t} - \rho \beta_l l_{i,t-1} + \beta_k k_{i,t} - \rho \beta_k k_{i,t-1} \\
& + \beta_{ll} l_{i,t}^2 - \rho \beta_{ll} l_{i,t-1}^2 + \beta_{lk} l_{i,t} k_{i,t} - \rho \beta_{lk} l_{i,t-1} k_{i,t-1} + \beta_{kk} k_{i,t}^2 - \rho \beta_{kk} k_{i,t-1}^2 \\
& + (1 - \rho) \varphi_i + u_{i,t} + (e_{i,t} - \rho e_{i,t-1}).
\end{aligned} \tag{9}
$$

This equation has non-linear restrictions on $-\rho\beta_l$, $-\rho\beta_k$, $-\rho\beta_{ll}$, $-\rho\beta_{lk}$ and $-\rho\beta_{kk}$. Blundell and Bond (2000) suggest to impose these restrictions *ex post* using a minimum distance procedure, after obtaining a consistent linear estimator. In the present study, we estimate the equation while imposing the non-linear restrictions using the Gauss-Newton iteration method. The procedure and the associated standard errors are described in the appendix. This procedure estimates fewer parameters than the procedure used by Blundell and Bond (2000) in their first step. We expect that our procedure will lead to smaller standard errors, and hence a more precise estimation.

Remember that, in case of the first-differenced panel estimation, having lagged dependent variable $y_{i,t-1}$ as a regressor causes yet another endogenous problem.[20] The second or longer lag of $y_{i,t}$ can be used as instrumental variables.

5.1.2 Olley and Pakes (OP) and Levisohn and Petrin (LP)

In parallel with the GMM applications to estimation of production functions, another class of estimation methods has been developed. Mainly initiated by Olley and Pakes (1996), these methods use some proxy variables, instead of instrumental variables, to control for the endogeneity problems between factor inputs and unobserved productivity. To illustrate, we assume a Cobb-Douglas

---

[20] This is because $E\left[(y_{i,t-1} - y_{i,t-2})(u_{i,t} - u_{i,t-1})\right] \neq 0$ due to $E(y_{i,t-1} u_{i,t-1}) \neq 0$. The bias is negative.

function

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \omega_{i,t} + e_{i,t}, \tag{10}$$

where $\omega_{i,t}$ is productivity and $e_{i,t}$ is a measurement error. The intercept β_0 is assumed to be included in $\omega_{i,t}$. We assume a translog function in the subsequent estimations.

Olley and Pakes (1996) propose investment data as a proxy for unobserved productivity because, even though investment is decided before current productivity $\omega_{i,t}$ is known to the firm, the firm's investment decision relies on its expectation on $E[\omega_{i,t}|\omega_{i,t-1}]$ and therefore correlates with productivity under the assumption that the innovation $\xi_{i,t}$ in the first-order Markov process $\omega_{i,t} = E[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t}$ is uncorrelated with the firm's information set at time $t-1$. Under this assumption, investment can be expressed as a function of productivity and capital as $i_{i,t} = i_t(\omega_{i,t}, k_{i,t})$, where $i_{i,t}$ denotes investment and $\omega_{i,t}$ unobserved productivity. Olley and Pakes then invert the function, $\omega_{i,t} = i_t^{-1}(k_{i,t}, i_{i,t})$, and use the inverted function as a proxy for productivity.

This approach works only if investment is a strictly monotonic function of productivity. However, empirical studies are often concerned with violation of this assumption due to lumpy investment (frequent observations with zero value). Levisohn and Petrin (2003) propose intermediate inputs (raw materials) as an alternative to investment as a proxy. Instead of $i_t(\omega_{i,t}, k_{i,t})$, they assume $m_{i,t} = m_t(\omega_{i,t}, k_{i,t})$, where $m_{i,t}$ is intermediate inputs, and express productivity as $\omega_{i,t} = m_t^{-1}(k_{i,t}, m_{i,t})$.

Both approaches consist of two stages, where the coefficient on labor, β_l, is identified in the first stage:

$$y_{i,t} = \beta_l l_{i,t} + \Phi(k_{i,t}, i_{i,t}) + e_{i,t}, \tag{11}$$

where $\Phi(k_{i,t}, i_{i,t}) = \beta_k k_{i,t} + i_t^{-1}(k_{i,t}, i_{i,t})$ (for Levisohn and Petrin's approach, $\Phi(k_{i,t}, i_{i,t})$ is replaced by $\Phi(k_{i,t}, m_{i,t})$ and $i_t^{-1}(k_{i,t}, i_{i,t})$ is replaced by $m_t^{-1}(k_{i,t}, m_{i,t})$). $\Phi(k_{i,t}, i_{i,t})$ can be treated nonparametrically or substituted by a polynomial approximation. Identification of β_l is then successfully done.

Given the estimate of $\Phi(k_{i,t}, i_{i,t})$, β_k is identified in the second stage. Since

$$\Phi(k_{i,t}, i_{i,t}) = \beta_k^* k_{i,t} + \omega_{i,t}, \tag{12}$$

we can obtain $\omega_{i,t}$ for every value of β_k^*; namely $\omega_{i,t}(\beta_k^*)$. According to Ackerberg et al.(2006), given the assumption that $\omega_{i,t}$ follows a first-order Markov

process

$$\omega_{i,t}\left(\beta_k^*\right) = E\left[\omega_{i,t}|\omega_{i,t-1}\right] + \xi_{i,t} = \Psi\left[\omega_{i,t-1}\left(\beta_k^*\right)\right] + \xi_{i,t}, \tag{13}$$

where $\xi_{i,t}$ is pure innovation, we choose a value of β_k^* that satisfies the moment condition

$$E\left[\xi_{i,t}k_{i,t}\right] = 0 \tag{14}$$

using a GMM method. Another way to identify β_k^*, proposed by Levisohn and Petrin (2003) and Petrin et al.(2004), is minimization of $\left(e_{i,t} + \xi_{i,t}\right)^2$:

$$\min_{\beta_k^*}\left(e_{i,t} + \xi_{i,t}\right)^2 = \min_{\beta_k^*}\left\{y_{i,t} - \hat{\beta}_l l_{i,t} - \beta_k^* k_{i,t} - \Psi\left[\omega_{i,t-1}\left(\beta_k^*\right)\right]\right\}^2. \tag{15}$$

In both cases, $\Psi\left[\omega_{i,t-1}\right]$ can be treated nonparametrically or substituted by a polynomial approximation.

5.1.3 Ackerberg, Caves, and Frazer's (ACF) approach

Levisohn and Petrin (2003) provide an alternative approach to the one proposed by Olley and Pakes (1996) to overcome some of its inherent limitations. However, Levisohn and Petrin's approach has been criticized by Ackerberg, Caves, and Frazer (2006) because identification of β_l in the first stage can be difficult. By assumption, labor $l_{i,t}$ and intermediate inputs $m_{i,t}$ are both flexible and chosen at the same time, which leads to collinearity of $l_{i,t}$ with the other terms.

Thus, Ackerberg et al.(2006) propose yet another approach. Their proxy for productivity $\omega_{i,t}$ consists of $m_{i,t}$, $k_{i,t}$ and $l_{i,t}$, namely $\omega_{i,t} = f_t^{-1}\left(m_{i,t}, k_{i,t}, l_{i,t}\right)$, and they assume productivity to follow a first-order Markov process. Although their approach also consists of two stages, they do not estimate any coefficient in interest in the first stage but only separate measurement errors. In the second stage, they identify the coefficients on labor and capital using a GMM method with two moment conditions, $E\left[\xi_{i,t}k_{i,t}\right] = 0$ and $E\left[\xi_{i,t}l_{i,t-1}\right] = 0$.

More specifically, they estimate in the first stage

$$y_{i,t} = \Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right) + e_{i,t}, \tag{16}$$

where $\Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right) = \beta_l l_{i,t} + \beta_k k_{i,t} + f_t^{-1}\left(m_{i,t}, k_{i,t}, l_{i,t}\right)$. Given the estimate of $\Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right)$, we can obtain $\omega_{i,t}\left(\beta_l^*, \beta_k^*\right)$ through $\omega_{i,t} = \Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right) - \beta_l^* l_{i,t} - \beta_k^* k_{i,t}$ for every value of β_l^* and β_k^*. Given the first-order Markov as-

sumption on $\omega_{i,t}$ that

$$\omega_{i,t}\left(\beta_l^*, \beta_k^*\right) = E\left[\omega_{i,t}|\omega_{i,t-1}\right] + \xi_{i,t} = \Psi\left[\omega_{i,t-1}\left(\beta_l^*, \beta_k^*\right)\right] + \xi_{i,t}, \qquad (17)$$

we choose values of β_k^* and β_k^* that satisfy

$$E\left[\xi_{i,t} \begin{array}{c} l_{i,t-1} \\ k_{i,t} \end{array}\right] = \mathbf{0}. \qquad (18)$$

Again, $\Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right)$ and $\Psi\left[\omega_{i,t-1}\right]$ can be treated nonparametrically or substituted by a polynomial approximation.[21]

5.1.4 Linearized ACF approach

We consider a special case where we have a linear relationship between $\omega_{i,t}$ and $\omega_{i,t-1}$ so that $\Psi\left[\omega_{i,t-1}\right]$ is a linear function. In this case, the ACF approach is reduced to a linear model

$$y_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \omega_{i,t} + e_{i,t}, \qquad (19)$$

where $e_{i,t}$ is a measurement error that is not correlated with production factors. Productivity $\omega_{i,t}$ evolves according to an autoregressive process

$$\omega_{i,t} = \psi\omega_{i,t-1} + \xi_{i,t}, \qquad (20)$$

where ψ is an autoregressive coefficient and $\xi_{i,t}$ is pure innovation. The reduced form is

$$\begin{aligned} y_{i,t} &= \psi y_{i,t-1} + \beta_l l_{i,t} - \psi\beta_l l_{i,t-1} + \beta_k k_{i,t} - \psi\beta_k k_{i,t-1} \\ &\quad + \xi_{i,t} + \left(e_{i,t} - \psi e_{i,t-1}\right). \end{aligned} \qquad (21)$$

This can be estimated by a linearized regression model using the Gauss-Newton iteration method. Because of the orthogonal conditions expressed by Eq.(18), $l_{i,t}$ is instrumented by $l_{i,t-1}$. This reduced form is very similar to Eq.(9) apart from the difference in type of production function. However, a critical difference

---

[21] Wooldridge (2009) suggests that Levisohn and Petrin' and Ackerberg, Caves, and Frazer's approaches are readily practicable in a one-step GMM framework. However, because his estimation approach is based on Levisohn and Petrin's assumption of $\omega_{i,t} = m^{-1}\left(k_{i,t}, m_{i,t}\right)$, it is subject to Ackerberg, Caves, and Frazer's criticism. His approach is, thereby, not an alternative to the ACF approach.

is that Eq.(21) does not have the firm-specific productivity component λ_i (see below).

We estimate the translog production function using Ackerberg, Caves, and Frazer's (ACF) framework. Since we apply it to a translog function, there are five orthogonality conditions,

$$E\left[\ \xi_{i,t}\ \left(\ l_{i,t-1}\quad k_{i,t}\quad l_{i,t-1}^2\quad l_{i,t-1}k_{i,t}\quad k_{i,t}^2\ \right)'\ \right]=\mathbf{0},\qquad(22)$$

to solve for the five coefficients β_l, β_k, β_{ll}, β_{lk}, and β_{kk}.

5.1.5   Relationship to the dynamic panel models

These new approaches try to tackle the three problems mentioned earlier, i.e., (i) individual effects, (ii) endogeneity of productivity components with regressors, and (iii) autocorrelaton in productivity components, in different ways from the fixed-effect approach and the GMM approach. While the fixed-effect and the GMM approaches try to solve the endogeneity problem through instrumental variables, these new approaches make use of proxies for productivity. They also address autocorrelation in productivity through a first-order Markov process of $\omega_{i,t}$. Related to this issue, a rather outstanding feature of the new approaches is that they do not assume time-invariant individual effects (i.e., λ_i in Eq.(7)). Individual differences in productivity are expressed by $\omega_{i,t}$, and they may differ among individuals. However, they develop according to a first-order Markov process and thus have no long-run fixed mean as λ_i.

5.2   Comparison between different estimators

5.2.1   Panel model approach

In this section, we compare results from different estimation methods. Table 3 presents estimation results using OLS, within-group (FE), difference GMM, and system GMM. OLS and GMM include year and 3-digit sector dummies, whereas FE and difference GMM only include year dummies.[22] In the following, we set the 3-digit SNI code "385: Treatment and coating of metals; general mechanical engineering" as benchmark. This is the subsector that contains the largest number of firms.

---

[22] We have also estimated the model including interaction terms between the 2-digit sector dummies and the year dummies to capture sector-specific year effects. The results are very similar to those presented in Table 2.

Our focus is on the specifications including a potential autocorrelation in the unobserved productivity shock $v_{i,t}$ (see Eq.(8)). For reference, we also present estimation results for the specifications without autocorrelation, which is provided in the first column of each estimator in Table 3.

[Table 3: Estimation of the production function]

For easier interpretation of the results, we report output elasticities with respect to labor and capital at four representative points: the medians of $l_{i,t}$ for micro, small, medium and large firms, respectively. The corresponding values of $k_{i,t}$ are set to $\hat{h}(l_{i,t})$, estimated by a locally weighted regression of $k_{i,t}$ on $l_{i,t}$ and year and 3-digit sector dummies:

$$k_{i,t} = h(l_{i,t}) + \psi_j + \phi_t + \epsilon_{i,t}, \tag{23}$$

where $h(\cdot)$ is an unknown smooth function, ψ_j is sector-specific effects (3-digit level), and ϕ_t is time-specific effects. For estimation of $h(\cdot)$, we use the semi-parametric approach outlined by Yatchew (1998). The bandwidth is set to 0.3.

The first columns of Table 3 report the OLS and the within-group (FE) estimates. Standard errors presented in the parentheses are robust to misspecification such as heteroskedasticity. They are also clustered by firms. Homotheticity and constant returns to scale are rejected in all cases.[23] The monotonicity and the quasi-concavity conditions are satisfied for most observations. As previous studies show, the within-group (FE) method generates lower output elasticities. The autocorrelation parameter ρ is expected to be upward biased in OLS and downward biased in the fixed-effect model. Although the hypotheses of $\rho = 1$ and $\rho = 0$ in the respective models are rejected, the direction of the difference between the two estimates is as expected. The associated scale elasticity is slightly decreasing with firm size when we assume no autocorrelation, while it is slightly increasing when we assume autocorrelation.

As already discussed, the OLS and the fixed-effect estimators are supposed to be biased. We now try to reduce the bias by using the GMM approach. The difference GMM uses lags of the explanatory and the dependent variables in levels as instruments for contemporaneous differences. The system GMM additionally uses lags of the explanatory variables expressed in first differences

---

[23] The null hypothesis of the test of homotheticity is $\beta_{ll} + \beta_{lk} = 0$ and $\beta_{lk} + \beta_{kk} = 0$. The null hypothesis of the test of constant returns to scale is $\beta_{ll} + \beta_{lk} = 0$, $\beta_{lk} + \beta_{kk} = 0$ and $\beta_l + \beta_k = 1$.

as instruments for contemporaneous levels.

For the estimation of the model specification without autocorrelation in the residuals, the lag length is set to two and three. The first lags are not valid as they correlate with the differenced residuals. This issue is also related to the Arellano-Bond test, $AR(s)$, which is the test of the null hypothesis that the differenced residuals in period t and $t-s$ are uncorrelated. Under the assumption of serially uncorrelated residuals in levels[24], the differenced residuals in t and $t-1$ are correlated while the differenced residuals in t and $t-2$ are uncorrelated.

For the estimation of the model specification with autocorrelation in the residuals, the lag length of the variables used as instruments is set to two and three for the explanatory variables. For the dependent variable, however, it is set to three and four for the following reason. If there is a serially uncorrelated measurement error $e_{i,t}$ in Eq.(7), the error term in the differenced equation contains $e_{i,t-2}$. Then $y_{i,t-2}$ is not a valid instrument. Regarding the Arellano-Bond test on the null hypothesis of serially uncorrelated residuals in levels, the differenced residuals in t and $t-1$ and in t and $t-2$ are correlated while the differenced residuals in t and $t-3$ are uncorrelated.

The estimation results are presented under "difference GMM" and "system GMM" in Table 3. All estimates are obtained by a one-step estimation, and the presented standard errors are robust to heteroskedasticity. The estimated production functions using the difference GMM imply negative output elasticity with respect to capital. Monotonicity is hence violated in a majority of the observations. Griliches and Mairesse (1998) point out that the differencing procedure exacerbates the bias caused by measurement errors in the explanatory variables, and this can be a reason why the difference GMM gives an unsatisfactory result.

The system GMM appears to reduce the bias by simultaneously estimating the equations in levels. However, the rejections of the Hansen J test and the Difference-in-Hansen test of overidentification[25] show that the estimators are subject to misspecification or to invalidity of instruments. The Arellano-Bond test is satisfactory for the model specification with autocorrelation in the residuals, but not for the model specification without autocorrelation.

We also test other sets of instruments, for example longer lags. However,

---

[24] The moment conditions in the GMM estimation are only valid under the assumption of no serial correlation in the error term.

[25] Arellano and Bond (1991) show that the one-step Sargan test overrejects in the presence of heteroskedasticity. In contrast, the Hansen J statistics are robust to this problem, while it can be weakened by too many instruments.

none of the estimated models satisfy the specification tests simultaneously. We suspect that potential instability of the parameters over time or potential heterogeneity of the parameters across firms of different sizes causes the rejection of the specification tests.

5.2.2 Ackerberg, Caves, and Frazer's (ACF) approach

We now move to estimation using the ACF approach. Firstly, we estimate the linearized version of the ACF approach since our preliminary estimation using the generalized ACF approach shows that estimated coefficients vary largely depending on the initial values given to the coefficients. In contrast, the linearized version of the ACF approach provides only a single solution.

Because this approach estimates Eq.(21) using the instrumental variables $l_{i,t-1}$, $k_{i,t}$, $l_{i,t-1}^2$, $l_{i,t-1}k_{i,t}$, and $k_{i,t}^2$, we present the results under "IV (linearized ACF)" in Table 4. We impose the non-linear restrictions using the Gauss-Newton iteration method. The standard errors are clustered by firms. Most of the coefficients are strongly significant. Output elasticity with respect to labor is decreasing with firm size while output elasticity with respect to capital is slightly increasing with firm size. As a result, scale elasticity decreases from 1.15 to 0.97.

Then we estimate the model with the generalized ACF approach. We use a third-order polynomial approximation for $\Theta\left(m_{i,t}, k_{i,t}, l_{i,t}\right)$ in the first stage. For $\Psi\left[\omega_{i,t-1}\right]$ in the second stage,[26] we use either a third-order or a fourth-order polynomial approximation. The result of the former is presented under "ACF (3rd poly)" while the result of the latter is presented under "ACF (4th poly)" in Table 4. The initial values of the coefficients are set to the coefficients obtained from the linearized ACF approach. The standard errors are obtained by bootstrapping.

[Table 4: Estimation of the production function with the ACF approach]

As the table shows, "ACF (3rd poly)" and "ACF (4th poly)" provide similar results. The estimated production functions imply that scale elasticity is around

---

[26]We have also tested a third-order polynomial approximation for $\Psi\left[\cdot\right]$. The resulted estimates are almost identical with slightly larger standard errors. Since there is no loss in degrees of freedom, we choose fourth-order.

1.16 for micro firms and then decreases with firm size to 0.97 for large firms.[27] These elasticities are almost identical to the ones obtained from the linearized ACF approach.

Our results differ from the previous studies presented earlier, which report scale elasticity under 1. The study by Oh et al.(2009a), using similar but older Swedish micro data, report scale elasticity only marginally above 1. However, their dataset is dominated by larger firms. In contrast, our dataset covers in principle all firms active in Sweden with at least one employee, and is consequently dominated by micro and small firms. The mean employees in our dataset is 33 (median 6), compared to 159 (median 44) in the dataset used in Oh et al.(2009a). Since medium-sized firms in our study (50 – 249 employees), which correspond to their average firms, have scale elasticity closer to one, the difference between their and our results may largely be explained by the size composition of the firms in the datasets. Our result is consistent with Heshmati (2001), who explores the growth patterns of Swedish micro and small firms (1 to 100 employees) and finds a positive scale effect.

6 Effects of imperfect competition on the estimates of scale elasticity

Note that our measure of value added $Y_{i,t}$ is constructed by deflating the nominal value with the sector-average price index (at the 2-digit sector level). As individual data on firms' output price indices is seldom available, this is a standard practice in the empirical literature. However, Klette and Griliches (1996) and Mairesse and Jaumandreu (2005) point out that this may bring a serious bias to estimates of output elasticities, even when output price indices are available at a very detailed level of industry classification.

A bias arises when, for instance, product markets are differentiated and imperfectly competitive. For an intuitive understanding, let us assume the "true" production function

$$y_i = \alpha_0 + \alpha_k k_i + \alpha_l l_i, \tag{24}$$

---

[27] The estimated polynomial function $\Psi\left[\omega_{i,t-1}\right]$ is $\Psi\left[\omega_{i,t-1}\right] = 0.748\omega_{i,t-1} - 0.041\omega_{i,t-1}^2 - 0.007\omega_{i,t-1}^3$ for "ACF (3rd poly)", and $\Psi\left[\omega_{i,t-1}\right] = 0.705\omega_{i,t-1} - 0.041\omega_{i,t-1}^2 + 0.015\omega_{i,t-1}^3 + 0.005\omega_{i,t-1}^4$ for for "ACF (4th poly)".

and the demand function

$$y_i = y_S - \eta (p_i - p_S) + d_i, \tag{25}$$

where $-\eta$ denotes price elasticity ($\eta > 1$), y_i real output of a firm i, y_S real output of sector S, which the firm belongs to, p_i the output price of the firm i, p_S the sector-average price of sector S, and d_i a demand shift parameter. The demand function implies that the share of the firm's output is determined by the price deviation from the average of the sector. If we only deflate the nominal output $p_i + y_i$ with the sector-average price p_S, the resulting output index $\tilde{y}_i$ is

$$\tilde{y}_i = p_i + y_i - p_S. \tag{26}$$

Erasing the price index p_i and substituting the equation of y_i into the production function, we obtain

$$\begin{aligned} \tilde{y}_i &= \frac{\eta - 1}{\eta} (\alpha_0 + \alpha_k k_i + \alpha_l l_i) + \frac{1}{\eta} (y_S + d_i) \\ &= \tilde{\alpha}_0 + \tilde{\alpha}_k k_i + \tilde{\alpha}_l l_i + \alpha_1 y_S + \tilde{d}_i, \end{aligned} \tag{27}$$

where $\tilde{\alpha}_0 = \frac{\eta-1}{\eta}\alpha_0$, $\tilde{\alpha}_k = \frac{\eta-1}{\eta}\alpha_k$, $\tilde{\alpha}_l = \frac{\eta-1}{\eta}\alpha_l$, $\alpha_1 = \frac{1}{\eta}$, and $\tilde{d}_i = \frac{1}{\eta}d_i$. The derived scale elasticity is then $(\alpha_k + \alpha_l) \frac{\eta-1}{\eta}$, which is smaller than true scale elasticity $\alpha_k + \alpha_l$ because of $0 < \frac{\eta-1}{\eta} < 1$, resulting in a downward bias in the estimated scale elasticity unless $\eta = \infty$ (i.e., no market power for the firm on the demand side).

Klette and Griliches (1996) propose a solution to the problem of omitted individual price variables. If we add the sectoral real output y_S (which is available) in the regression, we derive the magnitude of the bias $\frac{\eta-1}{\eta}$ using an estimate of $\frac{1}{\eta}$. In this way, we can obtain the estimate of elasticities α_k and α_l. Again, when the product markets are competitive (i.e., $\eta = \infty$), the coefficient on y_S is not significantly different from zero.

This approach is still restrictive as η and hence the magnitude of bias are assumed to be constant. It may be expected that small and large firms face different market conditions: small firms operate under nearly perfect competition with prices being given while large firms have market power that enables them to raise prices without losing sales as much as small firms would. To allow price

elasticity to vary with firm size, we modify the equation as

$$\tilde{y}_i = \tilde{\alpha}_0 + \tilde{\alpha}_k k_i + \tilde{\alpha}_l l_i + (\alpha_1 + \alpha_2 l_i) y_S + \tilde{d}_i. \tag{28}$$

In this case, the inverse of price elasticity is expressed by $\frac{1}{\eta} = \alpha_1 + \alpha_2 l_i$, i.e., it depends on the number of employees (proxy for firm size).

We apply this approach to our translog production function. The estimation method is the linearized ACF approach with instrumental variables. The result is presented in Table 5. The presented coefficients are unadjusted (i.e., $\tilde{\alpha}_k$ and $\tilde{\alpha}_l$), while the elasticities presented below are adjusted one (i.e., α_k and α_l but remember that the estimation is based on a translog function).

[**Table 5: Estimation of the production function (taking into account market power and imperfect competition)**]

The result shows that the estimated price elasticity η is fairly large and increases with firm size. It satisfies the condition of $\eta > 1$. The estimated scale elasticities are fairly similar to the ones in Table 5. We conclude that the estimation of the production function presented in Table 4 is robust even when we take into account market power and imperfect competition.

7 Estimation of production functions by sector

This section examines whether the estimates of the production functions obtained so far at the aggregate level actually reflect production functions at the sectoral level. We estimate a translog production function for each 2-digit sector using the linearized ACF approach with the instrumental variable method.

Table 6 presents estimated output elasticities by sector. Output and scale elasticities are evaluated at the same points as in previous analysis (i.e., the median of the number of employees for each size category). For all sectors, scale elasticity decreases with firm size. The last rows show weighted average using the number of observations in each sector as weights. The weighted average of output elasticity varies from 1.15 to 0.93. This is in line with the result from the aggregate production function.

[**Table 6: Output and scale elasticities based on estimation of the production function by sector**]

8   Decomposition of productivity

In this section, we decompose labor productivity according to Eq.(4) based on
the translog production function estimated by the linearized ACF approach with
instrumental variables. We try to find an answer to the question of why larger
firms on average enjoy higher productivity. As we have shown, scale elasticity
is decreasing with firm size. This means that the answer to the question is not
economies of scale.

We substitute $k_{i,t}$ in Eq.(4) with $\hat{h}(l_{i,t})$ from Eq.(23) and decompose labor
productivity into the four components according to Eq.(4). The four compo-
nents are: (i) deviation from constant returns to scale; (ii) variation in capital
intensity; (iii) deviation from homotheticity; and (iv) individual firm effects.

Figure 5 presents each component at different firm sizes. The component
related to (i) deviation from constant returns to scale first increases but then
decreases with firm size. This is consistent with the increasing returns to scale
for small firms and decreasing returns to scale for large firms found in the
previous sections.

[**Figure 5: Decomposition of the estimated production function**]

The component related to (ii) variation in capital intensity can be inter-
preted as variations in the factor price ratio when the production function is
homothetic (see Section 4.1). However, as we have rejected the null hypothesis
of homotheticity, this interpretation is no longer valid. Capital intensity can
vary as a result of both a variable factor price ratio and a non-homothetic pro-
duction function. The effects of variable capital intensity on labor productivity
are mixed in the components related to (ii) variation in capital intensity and
(iii) deviation from homotheticity.

Figure 6 depicts the estimated capital intensity, i.e., $\hat{h}(l_{i,t})$. As can be seen,
capital intensity increases with firm size after the initial decline. To explore the
factors driving the increase in capital intensity, we investigate the factor price
ratio derived from the estimated production function. It is expressed by

$$\frac{w}{r} = \frac{\exp(k_{i,t})}{\exp(l_{i,t})} \cdot \frac{\beta_l + \beta_{ll}l_{i,t} + \beta_{lk}k_{i,t}}{\beta_k + \beta_{kk}k_{i,t} + \beta_{lk}l_{i,t}}, \tag{29}$$

where we use $\hat{h}(l_{i,t})$ in Eq.(23) for $k_{i,t}$. Figure 7 reports the factor price ratio
w/r along different values of $l_{i,t}$. It declines rapidly in the range $0 \leq l_{i,t} \leq 2$ (or

$1 \leq L_{i,t} \leq 7.38$), but then it is almost constant. This suggests that the increase in capital intensity is solely driven by non-homotheticity.

[**Figure 6: Capital intensity**]

[**Figure 7: Factor price ratio**]

The component (iv) individual firm effects presented in Figure 5 is the mean of $\omega_{i,t}$ in Eq.(32) conditional on firm size $l_{i,t}$ (i.e., $E\left[\omega_{i,t}|l_{i,t}\right]$). It is not quantitatively different from zero. Because the residual $\xi_{i,t}$ in Eq.(21) is determined so that $E\left[\xi_{i,t}|l_{i,t},l_{i,t-1},k_{i,t},k_{i,t-1},y_{i,t-1}\right] = 0$, the deviation of $E\left[\omega_{i,t}|l_{i,t}\right]$ from zero can be explained by white noise and correlations between $l_{i,t}$ and $l_{i,t-1}$, $k_{i,t}$, $k_{i,t-1}$, $y_{i,t-1}$.

What we have found so far is that small firms face increasing returns to scale while large firms face decreasing returns to scale. We have also found that capital intensity increases with firm size because of non-homothetic production technology, and not because of heterogenous factor price ratios. This result suggests that small firms and large firms operate with different technology.

Decreasing returns to scale while capital intensity is constant implies productive disadvantage for larger firms, because output Y does not increase as much as an increase in inputs, L and K. However, when capital intensity increases with firm size, this may compensate the productive disadvantage and, in some cases as in this study, lead to higher labor productivity for larger firms.

9 Conclusion

This study explores the factors explaining the positive correlation between labor productivity and firm size. Specifically, our question is whether the productive advantage for large firms is due to scale effects, higher capital intensity, or other reasons.

To investigate scale elasticity at different firm sizes, we estimate a translog production function based on firm-level panel data on Swedish mining and manufacturing firms from 1997 to 2006. The dataset is unique because it is extracted from a tax registry, and therefore covers almost all firms active in these industries in Sweden. It includes more than 28,000 firms and 177,000 observations, and a majority of the firms are micro and small firms. This is in contrast to previous studies, which mostly cover medium and large firms.

We apply different estimation approaches available in the productivity literature, from a fixed-effect (within-effect) model and a dynamic panel model with the generalized method of moments (GMM) to the approach proposed by Ackerberg et al.(2006). We also address the issue of market power possibly biasing the estimated scale elasticity. To tackle this problem, we use the approach proposed by Klette and Griliches (1996).

While both the difference GMM and the system GMM approaches suffer from specification errors, our modified version of the ACF approach provides a stable estimation on scale elasticity. The estimation results show that scale elasticity for micro firms is around 1.15 whereas scale elasticity for large firms is around 0.97. This indicates that micro and small firms operate under (locally) increasing returns to scale, which has not been found in previous studies, whereas medium and large firms face (locally) decreasing returns to scale. The results are robust even when we take into account potential biases due to market power and imperfect competition. We also estimate a translog production function at the 2-digit sectoral level. The results are consistent with those at the aggregate level.

Increasing returns to scale for small firms but decreasing returns to scale for large firms may appear contradictory to the fact that labor productivity is a positive function of firm size. Our investigation shows that production technology is expressed by a non-homothetic function and that this causes capital intensity to increase with firm size whereas the factor price ratio is homogeneous. An increase in capital intensity raises labor productivity more than a decrease in economies of scale lowers it. This explains the positive association between firm size and labor productivity.

Our results also suggest that small firms and large firms operate with different technology. From this study, however, we cannot determine whether large firms obtain their capital-intensive technology after entry through learning by doing or they already had it at entry. The former may suggest that growth policy should promote, for example, transfer of efficient technology to small firms so that they can grow. The latter may suggest, for example, that firms with potentially high efficiency should be identified and then supported to grow.

References

[1] Ackerberg, Daniel A.; Caves, Kevin; and Frazer, Garth (2006), "Structural identification of production functions," mimeo, UCLA Department of Economics.

[2] Acs, Zoltan J. and Audretsch, David B. (1989), "Small-Firm Entry in U.S. Manufacturing," *Economica*, 56(222), 255-65.

[3] Alvarez, Roberto and Crespi, Gustavo (2003), "Determinants of Technical Efficiency in Small Firms," *Small Business Economics*, 20(3), 233-44.

[4] Andersson, Linda and Karpaty, Patrik (2007), "Offshoring and Relative Labor Demand in Swedish Firms," Working Papers 2007:5, Örebro University.

[5] Andersson, Martin and Lööf, Hans (2009a), "Agglomeration and Productivity – evidence from firm-level data," Working Paper Series in Economics and Institutions of Innovation, No. 170, Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology.

[6] Andersson, Martin and Lööf, Hans (2009b), "Learning-by-Exporting Revisited: The Role of Intensity and Persistence," *Scandinavian Journal of Economics*, 111(4), 893-916.

[7] Andersson, Martin and Lööf, Hans (2010),"Imports, Productivity and Origin Markets: The Role of Knowledge-intensive Economies," *The World Economy*, 33(3), 458-81

[8] Andersson, Martin; Lööf, Hans; and Johansson, Sara (2008), "Productivity and International Trade: Firm Level Evidence from a Small Open Economy," *Review of World Economics*, 144(4), 774-801.

[9] Arellano, Manuel and Bond, Stephen (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277-97.

[10] Arellano, Manuel and Bover, Olympia (1995), "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68(1), 29-51.

[11] Audretsch, David B. (2005), "The Emergence of Entrepreneurship Policy," *International Handbook Series on Entrepreneurship*, Volume 1, Part 7, 473-511.

[12] Badunenko, Oleg; Fritsch, Michael; and Stephan, Andreas (2008), "What Drives the Productive Efficiency of a Firm? – the importance of industry, location, R&D, and size," Working Paper Series in Economics and Institutions of Innovation, No. 126, Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology.

[13] Bain, Joe S. (1956), *Barriers to New Competition : Their Character and Consequences in Manufacturing Industries*, Harvard University Press.

[14] Baumol, William J. (2004), "Education for Innovation: Entrepreneurial Breakthroughs vs. Corporate Incremental Improvements," NBER Working Papers 10578.

[15] Bigsten, Arne; Collier, Paul; Dercon, Stefan; Fafchamps, Marcel; Gauthier, Bernard; Gunning, Jan W.; Oduro, Abena; Oostendorp, Remco; Pattillo, Catherine; Söderbom, Måns; Teal, Francis; and Zeufack, Albert (2004), "Do African Manufacturing Firms Learn from Exporting?" *The Journal of Development Studies*, 40(3), 115-141.

[16] Bjuggren, Carl Magnus; Daunfeldt, Sven-Olov; and Johansson, Dan (2010), "Ownership and High-Growth Firms," Ratio Working Papers No.147, The Ratio Institute.

[17] Blundell, Richard and Bond, Stephen (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115-43.

[18] Blundell, Richard and Bond, Stephen (2000), "GMM Estimation with persistent panel data: an application to production functions," *Econometric Reviews*, 19(3), 321-40.

[19] Brouwer, Peter; de Kok, Jan; and Fris, Pieter (2005), "Can firm age account for productivity differences?" EIM SCALES-paper N200421, Zoetermeer, Netherlands.

[20] Carlsson, Bo (1992), "The Rise of Small Business: Causes and Consequences," in Adams, William J. (ed.), *Singular Europe, Economy and Pol-*

icy of the European Community After 1992, University of Michigan Press, 145-69.

[21] Castany, Laia; López-Bazo, Enrique; and Moreno, Rosina (2005), "Differences In Total Factor Productivity Across Firm Size – A Distributional Analysis," ERSA conference papers, European Regional Science Association.

[22] Cohen, Wesley M. (1995), "Empirical Studies of Innovative Activity," in Stoneman, P. (ed), *Handbook of the Economics of Innovation and Technological Change*, 182-264, Blackwell.

[23] Cohen, Wesley M. and Klepper, Steven (1996), "Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D," *The Review of Economics and Statistics*, 78(2), 232-43.

[24] Cohen, Wesley M. and Levin, Richard C. (1989), "Empirical studies of innovation and market structure," in Schmalensee, R. and Willig, R. (eds.), *Handbook of Industrial Organization*, vol. 2, chap. 18, 1059-1107, Elsevier.

[25] Davis, Steven J. and Haltiwanger, John C. (1992), "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 107(3), 819-63.

[26] Davis, Steven J. and Henrekson, Magnus (1997a), "Industrial Policy, Employer Size, and Economic Performance in Sweden," in Freeman, R. B.; Topel, R. and Swedenborg, B. (eds.), *The Welfare State in Transition*, University of Chicago Press.

[27] Davis, Steven J. and Henrekson, Magnus (1997b), "Explaining National Differences in the Size and Industrial Distribution of Employment," NBER Working Paper No. 6246.

[28] De Loecker, Jan (2010), "A Note on Detecting Learning by Exporting," NBER Working Papers No. 16548, National Bureau of Economic Research.

[29] Dunne, Timothy; Roberts, Mark J.; and Samuelson, Larry (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, 19(4), 495-515.

[30] Ekholm, Karolina and Katariina Hakkala (2005), "The Effect of Offshoring on Labor Demand: Evidence from Sweden", Working Paper Series 654, Research Institute of Industrial Economics

[31] Eliasson, Kent; Hansson, Pär; and Lindvert, Markus (2009), "Do Firms Learn by Exporting or Learn to Export? Evidence from Small and Medium-Sized Enterprises (SMEs) in Swedish Manufacturing," Umeå Economic Studies 793, Umeå University.

[32] European Council (2000), "The European Observatory for SMEs," sixth report, Luxembourg.

[33] Geroski, Paul A. (1995), *Market Structure, Corporate Performance and Innovative Activity*, Clarendon Press, Oxford.

[34] Geroski, Paul A. (1999), "The Growth of Firms in Theory and in Practice," CEPR Discussion Papers 2092, Center for Economic Policy Reseach.

[35] Globalisation Council (2009a), *The role of SMEs and entrepreneurship in a globalised economy*, Ministry of Education and Research and the Swedish Cabinet Office.

[36] Globalisation Council (2009b), *Utvecklingskraft och omställningsförmåga - En globaliserad svensk ekonomi* (Ability for development and structural changes), the last report from the Globalisation Council, the Ministry of Education and Research and the Swedish Cabinet Office.

[37] Graham, Daniel J. (2001), "Productivity Growth in British Manufacturing: Spatial Variation in the Role of Scale Economies, Technological Growth and Industrial Structure," *Applied Economics*, 33(6), 811-21.

[38] Hansen, Thorsten (2010), "Exports and Productivity: An Empirical Analysis of German and Austrian Firm-Level Performance," Munich Discussion Paper, No. 2010-22, Department of Economics, University of Munich.

[39] Henrekson, Magnus (2005), "Entrepreneurship: a weak link in the welfare state?" *Industrial and Corporate Change*, 14(3), 437-467.

[40] Henrekson, Magnus and Johansson, Dan (1999), "Institutional Effects on the Evolution of the Size Distribution of Firms," *Small Business Economics*, 12(1), 11-23.

[41] Henrekson, Magnus and Johansson, Dan (2009), "Competencies and Institutions Fostering High-growth Firms," *Foundations and Trends in Entrepreneurship*, 5(1), 1-80.

[42] Heshmati, Almas (2001), "On the Growth of Micro and Small Firms: Evidence from Sweden," *Small Business Economics*, 17(3), 213-28.

[43] Heyman, Fredrik; Sjöholm, Fredrik; and Tingvall, Patrik (2007), "Is there really a foreign ownership wage premium? Evidence from matched employer-employee data," *Journal of International Economics*, 73(2), 355-76.

[44] Jovanovic, Boyan (1982), "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649-70.

[45] Karpaty, Patrik (2004), "Are foreign owned firms more productive? Evidence from Swedish firm data," Working Papers 2004:6, Örebro University.

[46] Karpaty, Patrik and Lundberg, Lars (2004), "Foreign Direct Investment and Productivity Spillovers in Swedish Manufacturing," Working Papers 2004:2, Örebro University.

[47] Klette, Tor J. and Griliches, Zvi (1996), "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous," *Journal of Applied Econometrics*, 11(4), 343-61.

[48] Levinsohn, James and Petrin, Amil (2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317-41.

[49] Loveman, Gary and Sengenberger, Werner (1991), "The re-emergence of small-scale production: An international comparison," *Small Business Economics*, 3(1), 1-31.

[50] Lööf, Hans (2008), "The Dynamics of Firm Growth – a re-examination," Working Paper Series in Economics and Institutions of Innovation, No. 128, Centre of Excellence for Science and Innovation Studies (CESIS), Swedish Royal Institute of Technology.

[51] Mairesse, Jacques and Jaumandreu, Jordi (2005), "Panel-data Estimates of the Production Function and the Revenue Function: What Difference Does It Make?" *Scandinavian Journal of Economics*, 107(4), 651-72.

[52] Nyström, Kristina (2005), "Firm Maturity and Product Processes R&D in Swedish Manufacturing Firms," Working Paper Series in Economics and Institutions of Innovation, No. 39, Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology.

[53] Oh, Donghyun; Heshmati, Almas; and Lööf, Hans (2009a), "Technical Change and Total Factor Productivity Growth for Swedish Manufacturing and Service Industries," Working Paper Series in Economics and Institutions of Innovation, No. 93, Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology.

[54] Oh, Donghyun; Heshmati, Almas; and Lööf, Hans (2009b), "Total Factor Productivity of Korean Manufacturing Industries: Comparison of Competing Models with Firm-Level Data," Working Paper Series in Economics and Institutions of Innovation, No. 201, Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology.

[55] Olley, G. Steven and Pakes, Ariel (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263-97.

[56] Pavcnik, Nina (2002), "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *Review of Economic Studies*, 69, 245-76.

[57] Poldahl, Andreas (2006), "Domestic vs. International Spillovers: Evidence from Swedish Firm Level Data," *Journal of Industry, Competition and Trade*, 6(3), 277-94.

[58] Petrin, Amil; Poi, Brian P.; and Levinsohn, James (2004), "Production function estimation in Stata using inputs to control for unobservables," *Stata Journal*, 4(2), 113-23.

[59] Rizov, Marian and Walsh, Patrick P. (2009), "Productivity and Trade Orientation in UK Manufacturing," *Oxford Bulletin of Economics and Statistics*, 71(6), 821-49.

[60] Schumpeter, Joseph A. (1942), *Capitalism, socialism, and democracy*, New York: Harper and Brothers.

[61] Stenkula, Mikael (2006), "The European Size Distribution of Firms and Employment," Working Paper Series No. 683, Research Institute of Industrial Economics.

[62] Söderbom, Mans and Teal, Francis (2004), "Size and efficiency in African manufacturing firms: evidence from firm-level panel data," *Journal of Development Economics*, 73(1), 369-94.

[63] Taymaz, Erol (2005), "Are Small Firms Really Less Productive?," *Small Business Economics*, 25(5), 429-45.

[64] Van Leeuwen, George and Klomp, Luuk (2006), "On the contribution of innovation to multi-factor productivity growth," *Economics of Innovation and New Technology*, 15(4-5), 367-90.

[65] Vial, Virginie (2006), "New Estimates Of Total Factor Productivity Growth In Indonesian Manufacturing," *Bulletin of Indonesian Economic Studies*, 42(3), 357-69.

[66] Wooldridge, Jeffrey M. (2009), "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 104(3), 112-14.

[67] Yang, Chih-Hai and Chen, Ku-Hsieh (2009), "Are small firms less efficient?," *Small Business Economics*, 32(4), 375-95.

[68] Yatchew, Adonis (1998), "Nonparametric Regression Techniques in Economics," *Journal of Economic Literature*, 36(2), 669-721

[69] Yip, George S. (1982), *Barriers to Entry: Corporate Strategy Perspective*, The Free Press.

Appendix: GMM Estimation of the model specified with residual autocorrelation using a linearized iteration method

A nonlinear model

$$\begin{aligned} y_t &= \rho y_{t-1} + \beta x_t - \rho\beta x_{t-1} + \epsilon_t \\ &= f(\mathbf{x}, \theta) + \epsilon_t \end{aligned} \tag{30}$$

is estimated by GMM using some instruments $\mathbf{z}$. One approach for this nonlinear estimation is a linearized regression model using the Gauss-Newton iteration method. Given that an estimate of $\theta \equiv (\beta, \rho)$ is obtained with this approach, the question is whether the estimate of the covariance matrix obtained at convergence is the same as the theoretically correct GMM covariance matrix for a nonlinear model.

Linearized regression model

A Taylor first-order approximation of $y = f(\mathbf{x}, \theta) + \epsilon$ at some initial value θ^0 is

$$
\begin{aligned}
f(\mathbf{x}, \theta) &\approx f(\mathbf{x}, \theta^0) + \sum_k \frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0} (\theta_k - \theta_k^0) \\
&= \left[f(\mathbf{x}, \theta^0) - \sum_k \theta_k^0 \left(\frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0} \right) \right] + \sum_k \theta_k \left(\frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0} \right). \quad (31)
\end{aligned}
$$

Putting it into the equation $y = f(\mathbf{x}, \theta) + \epsilon$ and placing the known terms on the left-hand size and the unknown terms on the right-hand side,[28]

$$
y - \left[f(\mathbf{x}, \theta^0) - \sum_k \theta_k^0 \left(\frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0} \right) \right] = \sum_k \theta_k \left(\frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0} \right) + \epsilon^0. \quad (32)
$$

With a value of θ_k^0 in hand, we regress the left-hand side on $\frac{f(\mathbf{x}, \theta^0)}{\partial \theta_k^0}$ to obtain an estimate of θ_k.

Our estimation strategy is to apply this approach to our model, estimate the linearized equation by GMM, and iterate the process until it converges. The left-hand side of Eq.(32) is equal to

$$
\begin{aligned}
& y_t - \left[\left(\rho^0 y_{t-1} + \beta^0 x_t - \rho^0 \beta^0 x_{t-1} \right) - \beta^0 \left(x_t - \rho^0 x_{t-1} \right) - \rho^0 \left(y_{t-1} - \beta^0 x_{t-1} \right) \right] \\
= \quad & y_t - \rho^0 \beta^0 x_{t-1}
\end{aligned} \quad (33)
$$

---

[28] ϵ^0 differs from ϵ except at convergence. It contains both the true disturbance ϵ and the error in the first-order Taylor approximation.

and the regressors $\frac{f(\mathbf{x},\theta^0)}{\partial \theta_k^0}$ are equal to

$$\frac{f\left(\mathbf{x},\theta^0\right)}{\partial\beta^0} \;=\; x_t - \rho^0 x_{t-1} \tag{34}$$

$$\frac{f\left(\mathbf{x},\theta^0\right)}{\partial\rho^0} \;=\; y_{t-1} - \beta^0 x_{t-1}. \tag{35}$$

That is, $y_t - \rho^0\beta^0 x_{t-1}$ is the dependent variable, while $x_t - \rho^0 x_{t-1}$ and $y_{t-1} - \beta^0 x_{t-1}$ are the two regressors. Given some initial values $\theta^0 = \left(\beta^0,\rho^0\right)$, we run a regression using GMM with the original set of instruments $\mathbf{z}$ to obtain an estimate of $\theta \equiv (\beta,\rho)$, and then set this value to a new θ^0 and run another regression to obtain a new estimate of θ. We iterate regressions until the difference between θ^0 and θ becomes very small. Note that we use the original set of instruments $\mathbf{z}$ instead of lags of the transformed regressors $x_t - \rho^0 x_{t-1}$, $y_{t-1} - \beta^0 x_{t-1}$ and dependent variable $y_t - \rho^0\beta^0 x_{t-1}$ as instruments.

The question here is if the covariance matrix estimated in the last regression of the iteration can be interpreted as the true covariance matrix for θ in the original nonlinear model. To answer this question, we will first describe the asymptotic variance of a GMM estimator for the original nonlinear model. We will thereafter show that it is the same as the one obtained in the last regression of the linearized model.

GMM covariance matrix

In general, the asymptotic variance of a GMM estimator $\tilde{\theta} \equiv \left(\tilde{\beta},\tilde{\rho}\right)$ is expressed as

$$\widehat{Avar\left(\tilde{\theta}\right)} = N^{-1}\left(\hat{\mathbf{G}}'\hat{\mathbf{\Lambda}}^{-1}\hat{\mathbf{G}}\right)^{-1}, \tag{36}$$

where

$$\hat{\mathbf{G}} \;\equiv\; N^{-1}\sum_i \boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\tilde{\theta}\right), \tag{37}$$

$$\hat{\mathbf{\Lambda}} \;\equiv\; N^{-1}\sum_i \mathbf{g}_i\left(\tilde{\theta}\right)\mathbf{g}_i\left(\tilde{\theta}\right)', \tag{38}$$

$$\mathbf{g}_i\left(\tilde{\theta}\right) \;\equiv\; \mathbf{g}\left(\mathbf{w}_i,\tilde{\theta}\right). \tag{39}$$

Applying this to our nonlinear model, $\mathbf{g}_i\left(\tilde{\theta}\right)$ and $\boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\tilde{\theta}\right)$ are expressed as

$$
\begin{aligned}
\mathbf{g}_i\left(\tilde{\theta}\right) &= \mathbf{z}_i' \tilde{\epsilon}_t \\
&= \mathbf{z}_i'\left(y_t - \tilde{\rho} y_{t-1} - \tilde{\beta} x_t + \tilde{\rho}\tilde{\beta} x_{t-1}\right)
\end{aligned}
\tag{40}
$$

and

$$
\begin{aligned}
\boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\tilde{\theta}\right) &= \left[\ \dfrac{\partial \mathbf{g}_i\left(\tilde{\theta}\right)}{\partial \tilde{\beta}} \quad \dfrac{\partial \mathbf{g}_i\left(\tilde{\theta}\right)}{\partial \tilde{\rho}}\ \right] \\
&= \left[\ \mathbf{z}_i'\left(-x_t + \tilde{\rho} x_{t-1}\right) \quad \mathbf{z}_i'\left(-y_{t-1} + \tilde{\beta} x_{t-1}\right)\ \right].
\end{aligned}
\tag{41}
$$

GMM covariance matrix from the linearized regression model

Because the dependent variable is $y_t - \rho^0\beta^0 x_{t-1}$ and the regressors are $x_t - \rho^0 x_{t-1}$ and $y_{t-1} - \beta^0 x_{t-1}$ in the linearized model, $\mathbf{g}_i\left(\hat{\theta}\right)$ and $\boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\hat{\theta}\right)$ for a GMM estimator $\hat{\theta} \equiv \left(\hat{\beta}, \hat{\rho}\right)$ given some $\theta^0 \equiv \left(\beta^0, \rho^0\right)$ are expressed as

$$
\begin{aligned}
\mathbf{g}_i\left(\hat{\theta}\right) &= \mathbf{z}_i' \hat{\epsilon}_t \\
&= \mathbf{z}_i'\left[y_t - \rho^0\beta^0 x_{t-1} - \hat{\beta}\left(x_t - \rho^0 x_{t-1}\right) - \hat{\rho}\left(y_{t-1} - \beta^0 x_{t-1}\right)\right]
\end{aligned}
\tag{42}
$$

and

$$
\begin{aligned}
\boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\hat{\theta}\right) &= \left[\ \dfrac{\partial \mathbf{g}_i\left(\tilde{\theta}\right)}{\partial \tilde{\beta}} \quad \dfrac{\partial \mathbf{g}_i\left(\tilde{\theta}\right)}{\partial \tilde{\rho}}\ \right] \\
&= \left[\ \mathbf{z}_i'\left(-x_t + \rho^0 x_{t-1}\right) \quad \mathbf{z}_i'\left(-y_{t-1} + \beta^0 x_{t-1}\right)\ \right].
\end{aligned}
\tag{43}
$$

Since $\hat{\theta} = \theta^0 = \theta^*$ (i.e., $\hat{\beta} = \beta^0 = \beta^*$ and $\hat{\rho} = \rho^0 = \rho^*$) at convergence,

$$
\mathbf{g}_i\left(\theta^*\right) = \mathbf{z}_i'\left(y_t - \rho^* y_{t-1} - \beta^* x_t + \rho^*\beta^* x_{t-1}\right)
\tag{44}
$$

and

$$
\boldsymbol{\nabla}_\theta \mathbf{g}_i\left(\theta^*\right) = \left[\ \mathbf{z}_i'\left(-x_t + \rho^* x_{t-1}\right) \quad \mathbf{z}_i'\left(-y_{t-1} + \beta^* x_{t-1}\right)\ \right].
\tag{45}
$$

So, given that θ^* exists and $\theta^* = \tilde{\theta}$, Eqs.(40) and (44) are identical. And Eqs.(41) and (45) are also identical.

	firm size				All
	micro	small	medium	large	
	(employees < 10)	(10 ≤ employees < 50)	(50 ≤ employees < 250)	(250 ≤ employees)	
number of firms in 1997	11376	4558	1283	349	17566
(proportion)	(0.65)	(0.26)	(0.07)	(0.02)	(1.00)
number of firms in 2006	12556	4694	1286	319	18855
(proportion)	(0.67)	(0.25)	(0.07)	(0.02)	(1.00)
share of value added	0.06	0.12	0.20	0.62	1.00
share of employment	0.06	0.15	0.23	0.53	1.00
share of capital stock	0.05	0.10	0.16	0.70	1.00
log of value-added per employee *	-1.542	-1.097	-0.965	-0.806	-1.423
(standard errors)	(0.008)	(0.009)	(0.011)	(0.019)	(0.009)
log of capital stock per employee *	-2.604	-2.134	-1.873	-1.573	-2.475
(standard errors)	(0.020)	(0.022)	(0.028)	(0.045)	(0.020)

Table 1: Summary statistics by firm size

* These data are the conditional means based on OLS regressions including dummy variables for sector, year and firm size

		distribution of firm size				
		micro	small	medium	large	
SNI-code	sector	(employees < 10)	(10 ≤ employees < 50)	(50 ≤ employees < 250)	(250 ≤ employees)	num. of firms
10-14	mining and quarrying	0.77	0.17	0.04	0.01	322
15-16	food products	0.68	0.25	0.05	0.02	2356
17-19	textiles	0.75	0.20	0.04	0.01	935
20	wood products	0.73	0.21	0.05	0.01	2634
21	pulp and paper	0.42	0.29	**0.15**	**0.13**	334
22	publishing and printing	0.76	0.19	0.05	0.01	3812
23	coke, petroleum and other fuels	0.53	0.15	**0.21**	**0.12**	34
24	chemicals	0.52	0.25	**0.17**	**0.05**	548
25	rubber and plastic products	0.62	0.29	0.09	0.01	1097
26	other non-metallic mineral products	0.65	0.22	0.10	0.03	640
27	basic metals	0.52	0.26	**0.15**	**0.07**	314
28	fabricated metal products	0.71	0.24	0.04	0.00	6460
29	machinery	0.63	0.26	0.09	0.02	3286
30	office machinery and computers	0.74	0.20	0.05	0.02	189
31	electrical machinery	0.59	0.30	0.09	0.02	878
32	communication equipment	0.58	0.28	0.08	0.06	368
33	medical and optical instruments	0.76	0.17	0.05	0.01	1174
34	motor vehicles	0.46	0.32	**0.15**	**0.07**	573
35	other transport equipments	0.74	0.18	0.05	0.02	579
36-37	furniture, recycling and else	0.71	0.23	0.06	0.01	1780

Table 2: Distribution of firm size in different sectors

Table 3: Estimation of the production function

	OLS		Within group (Fixed Effect)		difference GMM		system GMM	
lnL	0.8527 ***	0.7753 ***	0.7668 ***	0.6486 ***	1.685 ***	0.158	1.213 ***	0.335 ***
	(0.012)	(0.014)	(0.015)	(0.017)	(0.100)	(0.157)	(0.034)	(0.079)
lnK	0.2442 ***	0.212 ***	0.139 ***	0.0873 ***	-0.420 ***	0.020	0.183 ***	0.550 ***
	(0.008)	(0.009)	(0.010)	(0.011)	(0.071)	(0.084)	(0.018)	(0.049)
$(lnL)^2$	0.0228 ***	0.0277 ***	0.0147 ***	0.022 ***	-0.171 ***	0.0172	-0.035 ***	0.028
	(0.003)	(0.003)	(0.004)	(0.004)	(0.020)	(0.025)	(0.009)	(0.018)
lnL*lnK	-0.051 ***	-0.039 ***	-0.024 ***	-0.018 ***	0.129 ***	-0.039	-0.045 ***	-0.088 ***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.023)	(0.028)	(0.009)	(0.018)
$(lnK)^2$	0.0212 ***	0.0209 ***	0.0082 ***	0.0084 ***	-0.047 ***	-0.008	0.013 ***	0.036 ***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.009)	(0.010)	(0.003)	(0.006)
ρ		0.6165 ***		0.1727 ***		0.2695 ***		0.871 ***
		(0.006)		(0.008)		(0.035)		(0.014)
sector (3-digit) dummy	X	X	X	X	X	X	X	X
year dummy	X	X	X	X	X	X	X	X
Output elasticity for labor (L)								
micro firms (L = 3)	0.965 (0.004)	0.884 (0.006)	0.829 (0.007)	0.718 (0.009)	1.152 (0.043)	0.243 (0.100)	1.191 (0.016)	0.504 (0.044)
small firms (L = 17)	0.929 (0.004)	0.892 (0.006)	0.825 (0.009)	0.755 (0.011)	0.849 (0.059)	0.216 (0.088)	0.970 (0.016)	0.402 (0.048)
medium firms (L = 84)	0.912 (0.006)	0.911 (0.010)	0.830 (0.017)	0.794 (0.020)	0.529 (0.095)	0.203 (0.114)	0.781 (0.031)	0.336 (0.093)
large firms (L = 485)	0.886 (0.010)	0.926 (0.016)	0.831 (0.026)	0.834 (0.031)	0.199 (0.143)	0.183 (0.166)	0.566 (0.051)	0.249 (0.149)
Output elasticity for capital (K)								
micro firms (L = 3)	0.137 (0.003)	0.118 (0.004)	0.093 (0.004)	0.047 (0.004)	-0.165 (0.027)	-0.056 (0.024)	0.103 (0.007)	0.365 (0.024)
small firms (L = 17)	0.144 (0.003)	0.144 (0.004)	0.087 (0.005)	0.055 (0.005)	-0.151 (0.032)	-0.207 (0.080)	0.083 (0.010)	0.375 (0.025)
medium firms (L = 84)	0.137 (0.005)	0.155 (0.007)	0.078 (0.008)	0.056 (0.008)	-0.108 (0.040)	-0.343 (0.154)	0.056 (0.019)	0.361 (0.038)
large firms (L = 485)	0.136 (0.008)	0.174 (0.011)	0.069 (0.012)	0.060 (0.013)	-0.076 (0.054)	-0.494 (0.236)	0.030 (0.029)	0.358 (0.058)
Scale elasticity (the sum of the above two)								
micro firms (L = 3)	1.101 (0.004)	1.002 (0.005)	0.921 (0.008)	0.766 (0.010)	0.987 (0.032)	0.187 (0.107)	1.294 (0.016)	0.869 (0.051)
small firms (L = 17)	1.073 (0.002)	1.036 (0.003)	0.913 (0.009)	0.809 (0.011)	0.699 (0.059)	0.009 (0.143)	1.053 (0.009)	0.777 (0.044)
medium firms (L = 84)	1.049 (0.004)	1.066 (0.005)	0.907 (0.016)	0.850 (0.019)	0.421 (0.106)	-0.140 (0.218)	0.837 (0.024)	0.697 (0.087)
large firms (L = 485)	1.021 (0.006)	1.100 (0.009)	0.900 (0.026)	0.894 (0.030)	0.123 (0.161)	-0.311 (0.318)	0.597 (0.043)	0.607 (0.145)
Test of homotheticity (p-value)	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.231
Test of constant returns to scale (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Monotonicity (proportion of obs. that satisfy)	0.993	0.989	0.999	0.989	0.070	0.269	0.992	0.999
Quasi-concavity (proportion of obs. that satisfy)	0.995	0.996	0.995	0.996	0.881	0.993	0.995	0.995
Arellano-Bond test for AR(1) (p-value)					0.000	0.000	0.000	0.000
Arellano-Bond test for AR(2) (p-value)					0.000	0.000	0.000	0.000
Arellano-Bond test for AR(3) (p-value)					0.759		0.980	
Hansen J test (p-value)					0.000		0.000	
Difference-in-Hansen test (p-value)					0.000	0.000	0.000	0.000

	IV (linearized ACF)	ACF (3rd poly)	ACF (4th poly)
lnL	1.040 ***	0.990 ***	1.050 ***
	(0.021)	(0.123)	(0.173)
lnK	0.164 ***	0.211 **	0.169 **
	(0.014)	(0.085)	(0.126)
(lnL)^2	-0.002	0.011	-0.006
	(0.005)	(0.036)	(0.051)
lnL*lnK	-0.039 ***	-0.057	-0.035
	(0.007)	(0.047)	(0.070)
(lnK)^2	0.023 ***	0.028 **	0.022 **
	(0.003)	(0.013)	(0.021)
ρ	0.599 ***		
	(0.006)		
sector (3-digit) dummy	X	X	X
year dummy	X	X	X
Output elasticity for labor (L)			
micro firms (L = 3)	1.083 (0.008)	1.085 (0.017)	1.079 (0.027)
small firms (L = 17)	0.989 (0.005)	0.994 (0.034)	0.977 (0.046)
medium firms (L = 84)	0.915 (0.009)	0.928 (0.068)	0.894 (0.088)
large firms (L = 485)	0.827 (0.016)	0.847 (0.099)	0.798 (0.124)
Output elasticity for capital (K)			
micro firms (L = 3)	0.066 (0.005)	0.081 (0.005)	0.077 (0.006)
small firms (L = 17)	0.100 (0.003)	0.105 (0.021)	0.114 (0.029)
medium firms (L = 84)	0.118 (0.006)	0.110 (0.051)	0.134 (0.068)
large firms (L = 485)	0.144 (0.011)	0.125 (0.079)	0.163 (0.104)
Scale elasticity (the sum of the above two)			
micro firms (L = 3)	1.149 (0.006)	1.165 (0.018)	1.156 (0.024)
small firms (L = 17)	1.089 (0.003)	1.099 (0.013)	1.091 (0.017)
medium firms (L = 84)	1.033 (0.005)	1.039 (0.019)	1.028 (0.021)
large firms (L = 485)	0.971 (0.010)	0.972 (0.023)	0.961 (0.024)
Test of homotheticity (p-value)	0.000	0.000	0.000
Test of constant returns to scale (p-value)	0.000	0.000	0.000
Monotonicity (proportion of obs. that satisfy)	0.933	0.928	0.957
Quasi-concavity (proportion of obs. that satisfy)	0.985	0.992	0.993
Underindentification test (p-value)	0.000		
Overidentification test (p-value)	0.000		

Table 4: Estimation of the production function with the ACF approach

	IV (linearlized ACF)
lnL	1.061 ***
	(0.029)
lnK	0.163 ***
	(0.014)
(lnL)^2	-0.002
	(0.005)
lnL*lnK	-0.039 ***
	(0.007)
(lnK)^2	0.023 ***
	(0.003)
ρ	0.599 ***
	(0.006)
α1	0.034 ***
	(0.009)
α2	-0.003
	(0.003)
sector (3-digit) dummy	X
year dummy	X
Price elasticity (η)	
micro firms (L = 3)	31.8
small firms (L = 17)	37.1
medium firms (L = 84)	43.9
large firms (L = 485)	54.8
Output elasticity for labor (L)	
micro firms (L = 3)	1.140 (0.023)
small firms (L = 17)	1.038 (0.022)
medium firms (L = 84)	0.957 (0.023)
large firms (L = 485)	0.863 (0.027)
Output elasticity for capital (K)	
micro firms (L = 3)	0.068 (0.005)
small firms (L = 17)	0.103 (0.003)
medium firms (L = 84)	0.121 (0.007)
large firms (L = 485)	0.147 (0.011)
Scale elasticity (the sum of the above two)	
micro firms (L = 3)	1.208 (0.023)
small firms (L = 17)	1.141 (0.021)
medium firms (L = 84)	1.078 (0.021)
large firms (L = 485)	1.011 (0.023)
Test of homotheticity (p-value)	0.000
Test of constant returns to scale (p-value)	0.000
Monotonicity (proportion of obs. that satisfy)	0.933
Quasi-concavity (proportion of obs. that satisfy)	0.829
Underindentification test (p-value)	0.000

Table 5: Estimation of the production function (market power and imperfect competition)

Table 6: Output and scale elasticities based on estimation of the production function by sector

| SNI-code	sector	Output elasticity for labor (L)				Output elasticity for capital (K)				Scale elasticity (the sum of the two)			
		micro	small	medium	large	micro	small	medium	large	micro	small	medium	large
10-14	mining and quarrying	1.502 (0.084)	1.426 (0.104)	1.441 (0.145)	1.415 (0.189)	-0.074 (0.039)	-0.109 (0.057)	-0.203 (0.092)	-0.275 (0.131)	1.428 (0.061)	1.316 (0.051)	1.238 (0.060)	1.140 (0.075)
15-16	food products	1.060 (0.018)	0.922 (0.017)	0.786 (0.033)	0.640 (0.052)	0.017 (0.010)	0.142 (0.011)	0.247 (0.022)	0.368 (0.034)	1.077 (0.014)	1.064 (0.008)	1.033 (0.015)	1.008 (0.026)
17-19	textiles	1.216 (0.030)	0.976 (0.027)	0.752 (0.053)	0.508 (0.087)	0.044 (0.014)	0.117 (0.019)	0.177 (0.035)	0.246 (0.054)	1.260 (0.023)	1.093 (0.014)	0.929 (0.033)	0.754 (0.057)
20	wood products	1.087 (0.019)	0.996 (0.020)	0.931 (0.035)	0.850 (0.054)	0.074 (0.009)	0.089 (0.012)	0.087 (0.022)	0.092 (0.033)	1.162 (0.016)	1.086 (0.010)	1.018 (0.020)	0.943 (0.034)
21	pulp and paper	0.974 (0.077)	0.807 (0.076)	0.621 (0.118)	0.432 (0.165)	0.108 (0.053)	0.221 (0.052)	0.335 (0.077)	0.455 (0.104)	1.082 (0.038)	1.029 (0.030)	0.956 (0.046)	0.888 (0.069)
22	publishing and printing	1.170 (0.018)	1.043 (0.015)	0.942 (0.032)	0.822 (0.054)	0.067 (0.008)	0.057 (0.011)	0.036 (0.020)	0.019 (0.031)	1.238 (0.015)	1.100 (0.009)	0.978 (0.022)	0.842 (0.039)
24	chemicals	1.189 (0.044)	1.022 (0.038)	0.872 (0.069)	0.705 (0.109)	-0.036 (0.026)	0.078 (0.023)	0.165 (0.040)	0.269 (0.063)	1.153 (0.031)	1.100 (0.020)	1.036 (0.033)	0.974 (0.054)
25	rubber and plastic products	1.170 (0.030)	0.994 (0.028)	0.854 (0.050)	0.689 (0.078)	0.021 (0.016)	0.070 (0.017)	0.092 (0.032)	0.128 (0.050)	1.190 (0.025)	1.064 (0.014)	0.946 (0.027)	0.817 (0.048)
26	other non-metallic mineral products	1.065 (0.035)	0.958 (0.029)	0.874 (0.048)	0.774 (0.074)	0.115 (0.019)	0.142 (0.020)	0.152 (0.035)	0.171 (0.054)	1.180 (0.028)	1.100 (0.014)	1.026 (0.024)	0.945 (0.044)
27	basic metals	1.170 (0.059)	1.045 (0.050)	0.972 (0.075)	0.870 (0.111)	0.084 (0.034)	0.110 (0.033)	0.098 (0.054)	0.102 (0.082)	1.254 (0.038)	1.155 (0.023)	1.069 (0.026)	0.973 (0.043)
28	fabricated metal products	1.068 (0.010)	0.956 (0.009)	0.870 (0.018)	0.766 (0.029)	0.054 (0.006)	0.102 (0.006)	0.128 (0.012)	0.165 (0.019)	1.122 (0.008)	1.058 (0.005)	0.997 (0.012)	0.931 (0.021)
29	machinery	1.073 (0.013)	0.990 (0.011)	0.933 (0.019)	0.861 (0.032)	0.054 (0.009)	0.089 (0.008)	0.100 (0.016)	0.123 (0.025)	1.127 (0.010)	1.079 (0.005)	1.034 (0.009)	0.984 (0.017)
30	office machinery and computers	1.014 (0.049)	0.976 (0.040)	0.946 (0.083)	0.911 (0.137)	0.099 (0.026)	0.086 (0.039)	0.072 (0.071)	0.058 (0.108)	1.113 (0.041)	1.062 (0.021)	1.018 (0.044)	0.969 (0.081)
31	electrical machinery	1.012 (0.027)	0.964 (0.018)	0.936 (0.032)	0.897 (0.055)	0.103 (0.014)	0.101 (0.016)	0.086 (0.029)	0.077 (0.046)	1.116 (0.022)	1.065 (0.011)	1.023 (0.020)	0.974 (0.037)
32	communication equipment	1.139 (0.068)	1.056 (0.047)	0.955 (0.071)	0.856 (0.118)	-0.044 (0.036)	-0.005 (0.035)	0.062 (0.053)	0.118 (0.081)	1.095 (0.051)	1.061 (0.025)	1.017 (0.041)	0.974 (0.078)
33	medical and optical instruments	1.116 (0.030)	1.010 (0.035)	0.928 (0.067)	0.830 (0.105)	0.072 (0.017)	0.105 (0.026)	0.119 (0.046)	0.142 (0.069)	1.188 (0.023)	1.115 (0.014)	1.047 (0.032)	0.973 (0.055)
34	motor vehicles	1.010 (0.031)	0.984 (0.020)	0.978 (0.033)	0.963 (0.055)	0.095 (0.019)	0.092 (0.015)	0.074 (0.024)	0.062 (0.038)	1.105 (0.021)	1.075 (0.011)	1.052 (0.013)	1.025 (0.025)
35	other transport equipments	1.031 (0.030)	0.995 (0.025)	0.987 (0.044)	0.965 (0.071)	0.103 (0.016)	0.084 (0.021)	0.050 (0.037)	0.021 (0.057)	1.134 (0.025)	1.079 (0.013)	1.037 (0.023)	0.986 (0.041)
36-37	furniture, recycling and else	1.041 (0.023)	0.972 (0.021)	0.917 (0.035)	0.853 (0.055)	0.117 (0.011)	0.111 (0.015)	0.100 (0.026)	0.091 (0.040)	1.158 (0.019)	1.083 (0.010)	1.017 (0.021)	0.944 (0.037)
	TOTAL	1.096 (0.006)	0.987 (0.005)	0.900 (0.009)	0.797 (0.015)	0.059 (0.003)	0.094 (0.004)	0.111 (0.006)	0.137 (0.010)	1.156 (0.004)	1.081 (0.003)	1.010 (0.005)	0.934 (0.010)
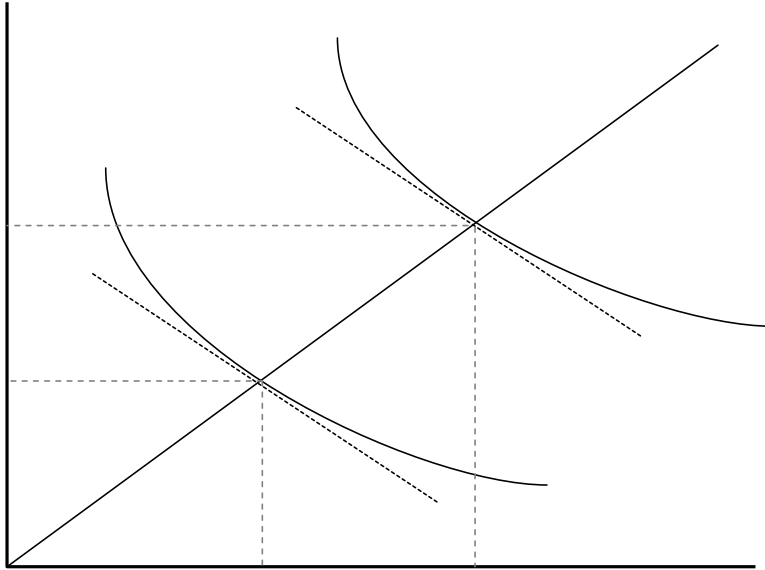
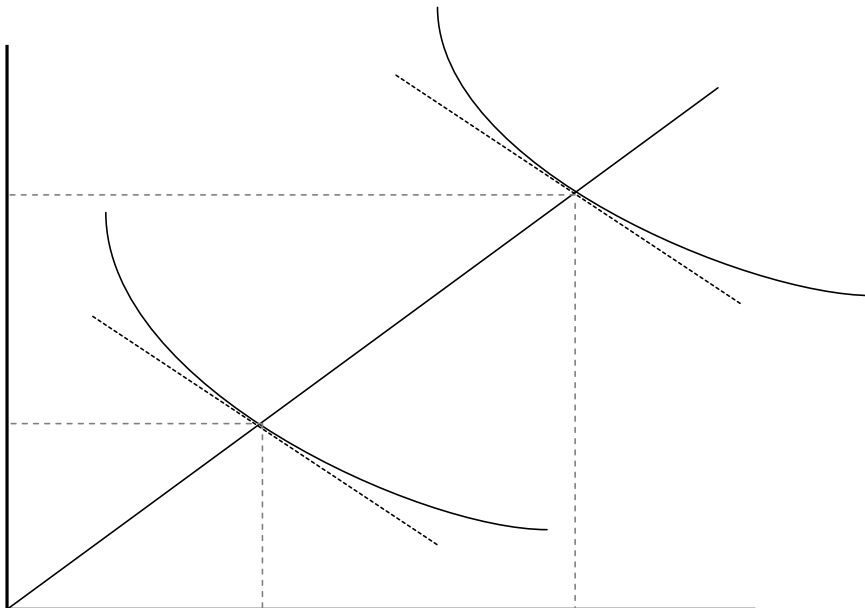Figure 1: Isoquants for output Y and 2Y under homotheticity and constant returns to scale

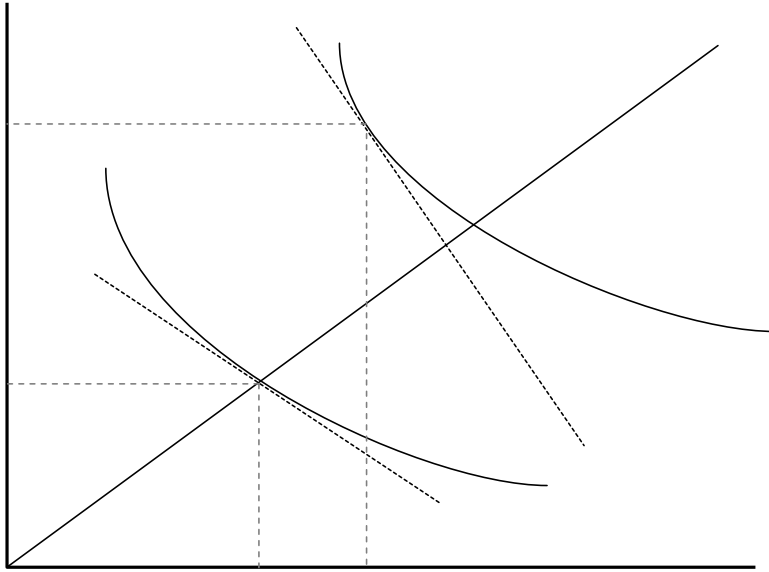Figure 2: Isoquants for output Y and 2Y under homotheticity but decreasing returns to scale

Figure 3: Isoquants for output Y and 2Y under homotheticity and constant returns to scale with heterogeneous factor price ratios
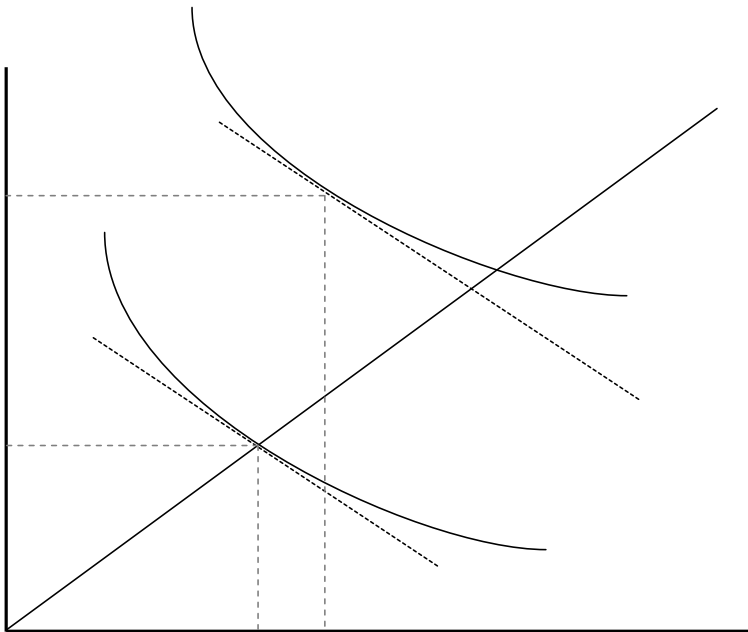

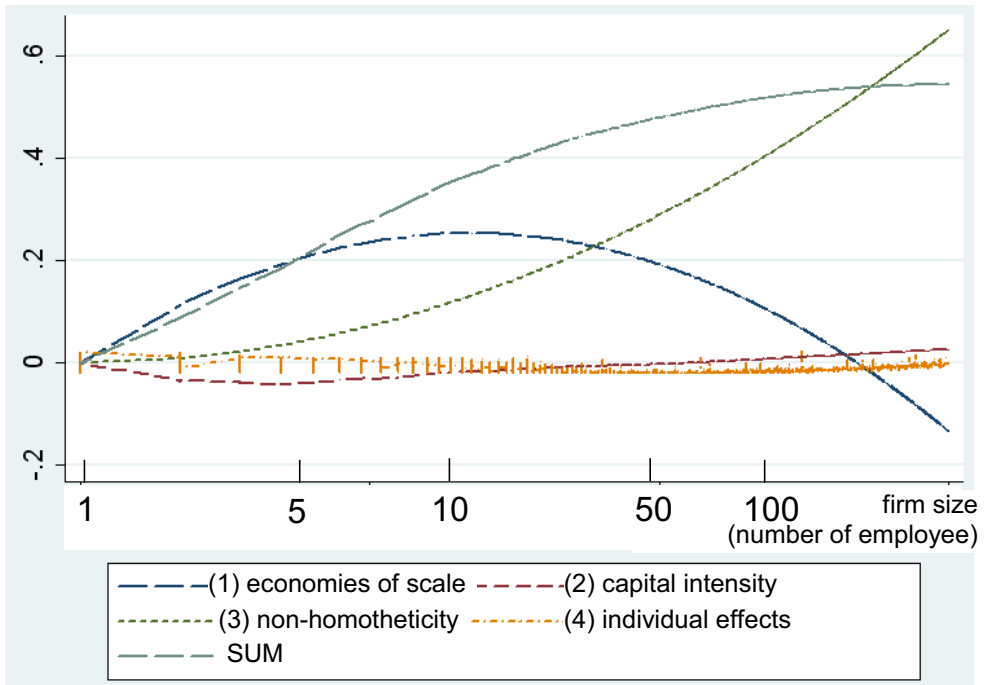Figure 4: Isoquants for output Y and 2Y under non-homotheticity

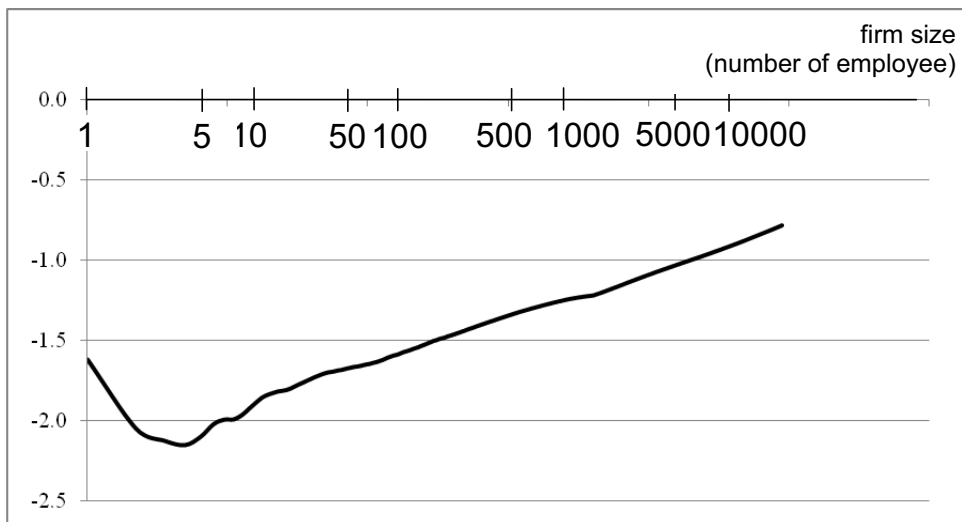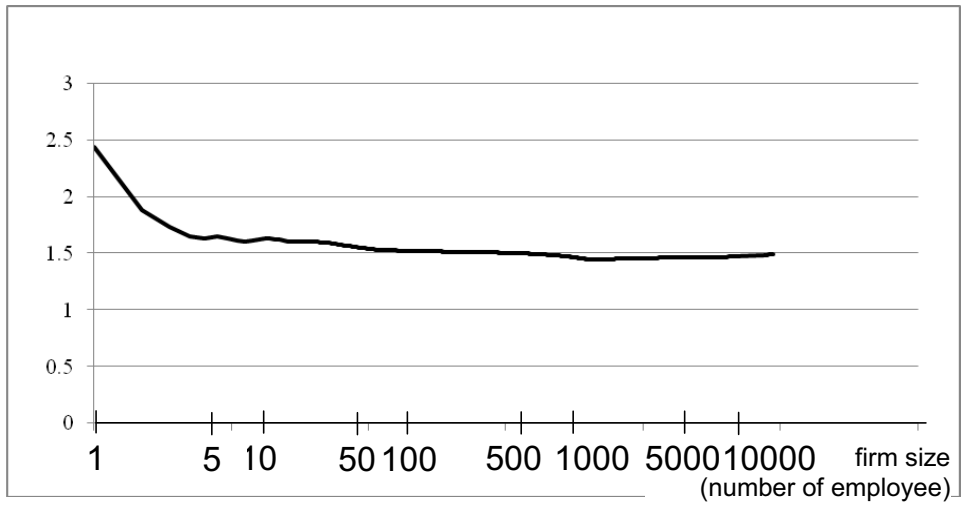Figure 5: Decomposition of labor productivity into four deferent components

Figure 6: Capital intensity

Figure 7: Factor price ratio (w/r)

Paper 4

Employment generation and productivity contribution of entrepreneurial firms compared to large incumbents

Yoshihiro Sato
University of Gothenburg

Abstract

Previous studies have reported that young and small so-called entrepreneurial firms have disproportionately large impacts on both employment generation and productivity growth. However, these positive impacts are conditional on the firms' survival. Many studies show that young and small firms have high mortality. In our study, we investigate the contributions of entrepreneurial firms to employment generation and productivity growth after taking the high mortality into account. We find that, although young and small firms are less likely to survive, they contribute more than other firms to both employment generation and aggregate productivity, simply because the survivors perform eminently.

1   Introduction

Generation of employment opportunities is a big policy issue. Dismissals of large numbers of workers by major employers make headline news and are regarded to have huge consequences both for the economy and for society at large. In Sweden, for example, employment was one of the most important factors behind the change in government from the Social Democrats to the center-right coalition after the general election in 2006 (Lindvall and Rueda, 2010). Also in other industrial countries, discussions of whether of not governments should support failed large companies to save jobs have been intensified especially since the world economic crises in 2008.

In recent decades, the role of small and young firms, so-called entrepreneurial firms, for employment creation has attracted attention in both research and policy debates, as a consequence of globalization. Confronted with cost competition and an increased degree of uncertainty, large-scale producers in high-cost

countries have lost competitiveness (Brock and Evans, 1989; Carlsson, 1992; and Audretsch, 2005). The comparative advantages in these countries have instead shifted toward knowledge-based economic activity, where scale economy is less important (Audretsch and Thurik, 2001). Technological progress, such as flexible automation and telecommunication, has facilitated and enhanced the structural shift.

In this new economic environment, small and new firms are regarded as driving force for economic development, in contrast to the previous view that they are associated with technological backwardness, managerial conservatism, and modest economic contributions (Storey, 2005). They enjoy lower labor costs and flexibility in specialization, compared to large and traditional capital-intensive firms (Acs and Audretsch, 1988 and Loveman and Sengenberger, 1991). Their entrepreneurial activities are vehicles for promoting new product innovation, discovering new markets and replacing inefficient incumbents, which provides conditions for evolution of the market structure and market performance (Geroski, 1995 and Audretsch, 2005). Moreover, Geroski (1995) points out that entry of new firms leads to erosion of high profits among incumbents and lower prices for consumers.

Thus, governments in industrial countries have become aware of the importance of entrepreneurial firms in economic growth and employment generation. They are therefore introducing a wide range of entrepreneurial policies, for example policies that facilitate access to finance and markets, that stimulate innovation and R&D, and provide information and advice for small businesses (to get a picture of different policy programs, see Table 3 in Storey, 2005).

The magnitude of the importance of entrepreneurial firms has been a subject of discussion in academic research in the last couple of decades (Davis et al., 1996; Carree and Klomp, 1996). Recent studies on entrepreneurial firms show that young and small firms play a very important role in employment generation, productivity growth, and innovation (Wennekers and Thurik, 1999; Audretsch and Thurik, 2000; Fölster, 2000; Baumol, 2002, 2005; Acs et al., 2005; Lundström and Stevenson, 2005; and van Praag and Versloot, 2007). For instance, several studies report a negative relationship between firm size and employment growth, i.e., small firms have disproportionately high contributions to the creation of jobs (Davis and Haltiwanger, 1992; Baldwin and Picot, 1995; Konings, 1995; Davis et al., 1996; Hart and Oulton, 1996; Broersma and Gautier, 1997; Sutton, 1997; Biggs, 2002; Johansson, 2005; Calvo, 2006; Oliveira and Fortunato, 2006; Shaffer, 2006; and Fritsch, 2008). The negative relationship

also applies to the relationship between firm age and employment growth rate (Davis and Haltiwanger, 1992; Konings, 1995; Calvo, 2006; and Oliveira and Fortunato, 2006). Large contributions of entrepreneurial firms are also reported in terms of value added growth (Robbins et al., 2000; Carree, 2002; Rodríguez et al., 2003; de Kok et al., 2005; and Carree and Thurik, 2006), labor productivity growth (Robbins et al. 2000; Disney et al., 2003; de Kok et al., 2005; and Carree and Thurik, 2006), and total factor productivity (TFP) growth (Callejon and Segarra, 1999).

However, the positive impacts of entrepreneurial firms reported in these previous studies are conditional on their survival. Dunne et al.(1988) and Davis and Haltiwanger (1992) find that young firms, especially if they are small, have a high mortality rate compared to incumbent firms. This is also reflected in a high reallocation rate of workers in entrepreneurial firms (Burgess et al., 2000) as well as a highly volatile growth rate (Davis and Haltiwanger, 1992; Baldwin and Picot, 1995; Davis et al., 1996; Lever, 1996; Broersma and Gautier, 1997; and Picot and Dupuy, 1998). This may imply that the selection effect based on survival leads to a seemingly higher employment and productivity growth rates among young and small firms. It is therefore necessary to take into account their higher mortality when comparing the potentials of small and young firms with those of incumbent firms. Although some of the previous studies estimate the employment growth rates using the Heckman model, the estimated survival rates and the growth rates unconditional on firm survival are not explicitly presented.

Another concern about entrepreneurial firms is their low productivity. Power (1998), Jensen et al.(2001), and de Kok et al.(2005) report that small and young firms on average have low productivity levels. This may imply that, even if small and young firms have higher growth rates in productivity, their contributions to aggregate productivity growth are not necessarily large. To correctly compare the productivity contributions of firms of different size and age, we need to develop a better measure of productivity contribution than productivity growth rates.

In the present study, we investigate the relative contribution of firms of different size and age to employment generation and productivity growth while accounting for the difference in survival rates. As a measure of employment contribution, we adapt employment growth rate both conditional and unconditional on firm survival. As a measure of productivity contribution, we propose using an individual firm's contribution to the aggregate productivity change per

employee.

The study is based on register-based panel data for the Swedish mining and manufacturing industries for the period 1997–2006. The dataset covers all legally independent firms active within the industries. Our dataset is unique in that it is dominated by micro firms, unlike in previous studies. A limitation of the dataset is that it is not possible to distinguish exits due to closures or bankruptcy from exits due to acquisitions or mergers.

We find that, although young and small firms have lower survival rates, their employment growth rates conditional on survival are on average higher than those of old and large firms. As a result of these opposite effects, the unconditional employment growth rate is highest for the firms that are about five years old and have less than five employees. We also observe that young and small firms contribute the most per employee to aggregate productivity growth. Even old, large firms exhibit relatively large per-employee contributions to aggregate productivity. However, the magnitude is much less compared to entrepreneurial firms.

Our contributions to the literature are as follows. First, we provide evidence of the important role played by small and young firms in employment generation and productivity growth while taking their high mortality into account. Second, we propose a measure of productivity contributions per employee, that can be interpreted easily.

The paper is organized as follows. In Section 2, we briefly review previous studies, followed by a discussions on our measures of employment generation and productivity contribution in comparison to the measures used in previous studies. Section 3 describes our dataset for empirical analysis. The descriptive statistics is presented in Section 4. In Section 5, we describe the econometric models for empirical analysis and then present the results. Section 6 concludes the study.

2 Conceptual framework

2.1 Previous studies

Previous studies on entrepreneurial firms include both macro-oriented studies using aggregate or sectoral data and micro-oriented studies using firm or establishment data. As our current study is based on micro panel data, we briefly review recent studies that use firm or establishment data.

In the literature on the relationship between employment generation and firm size and age, the most common studies deal with Gibrat's law of proportionate growth. Gibrat (1931) argues that mean growth rates in firm size, measured by the number of employees, are the same for all firm sizes independent of past growth history. However, the results from major studies such as Mansfield (1962), Evans (1987a, 1987b), Hall (1987), Dunne and Hughes (1994), Mata (1994), Audretsch and Mahmood (1994), Konings (1995), Hart and Oulton (1996), Audretsch et al.(1999a, 1999b), Almus and Nerlinger (2000), Heshmati (2001), Goddard et al.(2002), Nurmi (2004), Calvo (2006), and Oliveira and Fortunato (2006), can be summarized as follows: Generally speaking, Gibrat's law fails to hold at least for small firms. These studies namely find that small firms create more jobs than large firms. The negative relationship also applies to the relationship between firm age and employment growth rate (Davis and Haltiwanger, 1992; Konings, 1995; Calvo, 2006; and Oliveira and Fortunato, 2006). It should also be mentioned that Haltiwanger et al.(2010) suggest that the high growth rate of small firms can be explained by the fact that they are dominated by young firms. They report that there is no systematic relationship between firm size and growth once firm age is controlled for. Taking these studies together, firm growth rate is decreasing in firm age among firms of the same size, whereas it is uncorrelated or negatively related with firm size among firms of the same age.

Micro-based studies on productivity are divided into two groups. One consists of studies on productivity level and its relationship with firm size and age. De Kok et al.(2005) find that the mean labor productivity of large firms is significantly higher than that of small firms. Power (1998), Jensen et al.(2001), and de Kok et al.(2005) report that labor productivity is on average higher for older firms than younger firms.[1] De Kok et al.(2005) find a similar positive relationship between total factor productivity (TFP) and firm age, and Castany et al.(2005) report a positive relationship with firm size. On the other hand, Disney et al.(2003) find a negative relationship between both labor and total factor productivity and firm age, and Nguyen and Lee (2002) report that there is no significant relationship between TFP and firm size.

The other group consists of studies on productivity growth and its relationship with firm size and age. A majority of the previous studies find that productivity growth rates, in terms of either labor productivity or TFP, are neg-

---

[1]The productivity difference, however, disappears when they control for labor quality and capital intensity.

atively correlated with firm age (Power, 1998; Jensen et al., 2001; Celikkol, 2003; Disney et al., 2003; and Huergo and Jaumandreu, 2004). Jensen et al.(2001) find that entering cohorts converges with incumbent firms in terms of productivity in 5–10 years after entry. Regarding firm size, de Kok et al.(2005) find a negative relationship with growth in labor productivity and TFP, while Castany et al.(2005) report that small and large firms have similar TFP growth rates.

The negative relationship between productivity growth and firm age, i.e., young firms grow more rapidly than large firms, can be a result of two factors. The first is "active learning." Ericson and Pakes (1992) argue that entrants' efforts to catch up with incumbents in order to be competitive drive up their productivity over time.

Jovanovic (1982), on the other hand, hypothesizes that entrants do not initially have precise information on their capability before entry, but learn their advantage relative to competitors by engaging in business. Firms that do not learn fast enough and are not able to catch up or keep up with the productivity growth rate in the market will be forced to exit due to market selection (Taymaz, 2002 and Balk and Hoogenboom-Spijker, 2003). As a result, the average performance of existing firms may increase over time even if the survivors' performance as such does not improve. This effect is often called the selection effect. Ericson and Pakes (1992) refers to this process as a "passive learning" process.[2]

This selection effect is hard to ignore. Previous studies report that young firms, especially if they are small, have a high mortality rate compared to incumbent firms (Dunne et al.; 1988, Davis and Haltiwanger, 1992). Barnes and Haskel (2000) report that entry and exit mainly take place in the lower productivity quintiles within an industry. This may imply that the selection effect based on survival leads to seemingly higher employment and productivity growth rates young and small firms.

Taken together, previous studies indicate both advantages and disadvantages of entrepreneurial firms. Small and young firms lag behind older or larger firms in terms of productivity. On the other hand, they grow faster in terms both of size and productivity. However, this is partly explained by high mortality of small and young firms, meaning that there are many unsuccessful firms among young firms. Furthermore, whereas productivity gaps may disappear between

---

[2]There is also a "vintage" effect. Because best-practice technologies are embodied in new capital, younger plants may perform better than earlier cohorts of entrants (Jensen et al., 2001).

old and young firms over time due to both the active learning process and the selection effects, we are not sure whether productivity gaps between small and large firms also disappear.

2.2 Employment growth rate

In this study, we investigate the relative contribution of firms of different size and age in employment creation and productivity growth while we taking the difference in survival rates into account. Our first interest is to estimate the expected level of employment growth rate unconditional on firm survival. Assuming that subsequent employment changes to time T depend on firm size and age at the initial point 0, the unconditional employment level is expressed by

$$E\left[L_{i,T}|L_{i,0}, a_{i,0}\right], \tag{1}$$

where $L_{i,t}$ and $a_{i,t}$ denote the number of employees and firm age for firm i at time t, respectively. Some of the firms go out of business before time T, and in that case $L_{i,T} = 0$. The unconditional level of employment can be rewritten using the survival rate of the firm as

$$E\left[L_{i,T}|L_{i,0}, a_{i,0}\right] = E\left[L_{i,T}|L_{i,T} > 0, L_{i,0}, a_{i,0}\right] \cdot \Pr\left[L_{i,T} > 0|L_{i,0}, a_{i,0}\right]. \tag{2}$$

That is, the unconditional expectation is a product of the conditional expectation and the survival rate. The unconditional growth rate of employment is then expressed by

$$\begin{aligned} g_i &= E\left[\tfrac{L_{i,T}-L_{i,0}}{L_{i,0}}|L_{i,0}, a_{i,0}\right] \\ &= E\left[\tfrac{L_{i,T}-L_{i,0}}{L_{i,0}}|L_{i,T} > 0, L_{i,0}, a_{i,0}\right] \cdot \Pr\left[L_{i,T} > 0|L_{i,0}, a_{i,0}\right] \\ &\quad - \left(1 - \Pr\left[L_{i,T} > 0|L_{i,0}, a_{i,0}\right]\right). \end{aligned} \tag{3}$$

The unconditional growth rate can be interpreted as the expected contribution of unit employment in 1997 to the employment change in the following years up to 2006, including both surviving and exiting firms.

The previous studies that investigate the validity of Gibrat's law focus on the firms that existed during the whole sample periods (Mansfield, 1962; Evans, 1987a, 1987b; Hall, 1987; Audretsch and Mahmood, 1994; Dunne and Hughes, 1994; Mata, 1994; Konings, 1995; Hart and Oulton, 1996; Audretsch et al.,

1999a, 1999b; Almus and Nerlinger, 2000; Heshmati, 2001; Goddard et al., 2002; Nurmi, 2004; Calvo, 2006; and Oliveira and Fortunato, 2006). The studies use empirical models such as

$$
\begin{aligned}
\tilde{g}_i &= E\left[\frac{L_{i,T}-L_{i,0}}{L_{i,0}}|L_{i,T}>0, L_{i,0}, a_{i,0}\right] \\
&= E\left[\ln L_{i,T} - \ln L_{i,0}|L_{i,0}>0, L_{i,0}, a_{i,0}\right] = f(L_{i,0}, a_{i,0}). \quad (4)
\end{aligned}
$$

This is actually the conditional growth rate of employment. None of these studies analyze the expected employment growth including unsuccessful firms.

Since this approach is potentially subject to sample selection bias[3], many of the studies apply a generalized Tobit model of the form

$$
\begin{aligned}
\tilde{g}_i &= f(L_{i,0}, a_{i,0}) \quad &\text{if } L_{i,T} > 0 \\
&= 0 \quad &\text{if } L_{l,T} = 0, \quad (5) \\
\Pr\left[L_{i,T}>0|Z_{i,0}\right] &= \Phi\left[h\left(Z_{i,0}\right)\right], \quad &(6)
\end{aligned}
$$

where $Z_{i,0}$ is a vector of explanatory variables for firm survival including $L_{i,0}$ and $a_{i,0}$, and $\Phi\left[\cdot\right]$ is the cumulative normal distribution function with unit variance. They use the estimation procedure proposed by Heckman (1979), and hence, they actually calculate the terms corresponding to $\Pr\left[L_{i,T}>0|L_{i,0}, a_{i,0}\right]$ in Eqs.(2) and (3). Nevertheless, since the focus of the studies is on the growth rate of continuing firms, neither estimated survival rates nor employment growth rates that include exiting firms are explicitly presented.

In this study, we estimate the survival rate $\Pr\left[L_{i,T}>0|L_{i,0}, a_{i,0}\right]$ and the conditional employment level $E\left[L_{i,T}|L_{i,T}>0, L_{i,0}, a_{i,0}\right]$ to derive the unconditional growth rate of employment g_i in Eq.(3). As g_i denotes for the employment change per employee at an individual firm, we can easily express the aggregate employment change as the sum of g_i times firm size $L_{i,0}$:

$$
\sum_i E\left[L_{i,T} - L_{i,0}|L_{i,0}, a_{i,0}\right] = \sum_i \left(g_i \cdot L_{i,0}\right). \quad (7)
$$

We can compare, for instance, the expected values of contribution to employment change for 50 small firms with 2 employees each and for a large firm with 100 employees.

---

[3]The sample selection bias arises because only those firms that survived during the whole studied period are included in the analysis. For example, if firm size is positively related with firm survival so that the disturbances in the survival function and in the growth function are positively correlated, the estimated growth rates will be biased upward.

2.3   Productivity contribution per employee

We are also interested in comparing contributions of firms of different size and age to aggregate productivity growth. For this purpose, we propose analyzing the per-employee contribution to aggregate productivity, which is defined as follows.

Because aggregate labor productivity is a weighted sum of individual labor productivity with the labor share as the weight, growth in aggregate labor productivity from time 0 to T can be expressed using individual labor productivity by

$$
\begin{aligned}
\frac{\sum_i Y_{i,T}}{\sum_i L_{i,T}} - \frac{\sum_i Y_{i,0}}{\sum_i L_{i,0}} &= \sum_i \left(\frac{Y_{i,T}}{L_{i,T}} \frac{L_{i,T}}{\sum_k L_{k,T}} \right) - \sum_i \left(\frac{Y_{i,0}}{L_{i,0}} \frac{L_{i,0}}{\sum_k L_{k,0}} \right) \\
&= \sum_i \left(LP_{i,T} \cdot s_{i,T} - LP_{i,0} \cdot s_{i,0} \right),
\end{aligned} \tag{8}
$$

where $LP_{i,t}$ denotes individual labor productivity, i.e., $LP_{i,t} = Y_{i,t}/L_{i,t}$, and $s_{i,t}$ the labor share of the firm, i.e., $s_{i,t} = L_{i,t}/\sum_k L_{k,t}$. An individual firm's contribution to aggregate productivity change is given by

$$
LP_{i,T} \cdot s_{i,T} - LP_{i,0} \cdot s_{i,0}. \tag{9}
$$

Comparison of this term between firms of different size is problematic since $s_{i,i}$ depends on firm size $L_{i,t}$. This term tends to be larger for larger firms.

We propose to analyze a standardized measure of this term by dividing it with initial firm size, which we call μ_i:

$$
\pi_i = \frac{LP_{i,T} \cdot s_{i,T} - LP_{i,0} \cdot s_{i,0}}{L_{i,0}}. \tag{10}
$$

π_i is independent of firm size and denote the per-employee contribution of the firm to aggregate productivity change. In a special case where $L_{i,0} = L_{i,T}$ and $\sum_k L_{k,0} = \sum_k L_{k,T}$, comparison of π_i yields the same order as comparison of the term $LP_{i,T} - LP_{i,0}$.

Previous studies that investigate the productivity development of entrepreneurial firms, such as Power (1998), Jensen et al.(2001), Celikkol (2003), Disney et al.(2003), and Huergo and Jaumandreu (2004), compare the growth rate of

productivity in the form of

$$\frac{\omega_{i,T} - \omega_{i,0}}{\omega_{i,0}} \simeq \ln \omega_{i,T} - \ln \omega_{i,0}, \qquad (11)$$

where $\omega_{i,t}$ is a productivity measure such as labor productivity or TFP. This is difficult to interpret because it can appear large if the initial level is low regardless of the volume of the contribution. The actual contribution of the firm to aggregate productivity is not clear from this productivity growth rate.

In contrast, our measure π_i directly shows individual firms' contribution per employee. It is also directly comparable between firms of different sizes. We can interpret π_i as a worker's average productivity performance at the firm. It can also be interpreted as an answer to the question: "If we have one employment in 1997, what is the expected contribution of this employment to the aggregate labor productivity growth between 1997 and 2006?"

3   Data

This study is based on firm-level panel data from Statistics Sweden's Structural Business Statistics. The original database contains detailed information on the income statements, balance sheets, and physical investment of all legally independent firms active in Sweden, including all forms of private and public firms except financial firms. Most of the data are obtained from registers at the Swedish national tax agency. All firms are classified according to NACE (Classification of Economic Activities in the European Community) with 5-digit classes. Our dataset is extracted from this database and covers manufacturing and mining firms in the period 1997–2006 (for more details about the database, see Chapter 3).

There are 101,515 firms and 556,930 observations in the database. Following major previous studies, all firms without employees (e.g., self-employed firms) in some or all years from 1997 to 2006 are dropped. Since this study follows the firms that existed in 1997, only these firms are chosen. As a result, our unbalanced panel dataset contains 17,566 firms and a total of 135,316 observations (the average number of observations per firm is 5.5).

A firm exit is defined when a firm no longer appears in the following year. There is a limitation in the dataset. It is not possible to distinguish exits due to closures or bankruptcy from exits due to acquisition or merger.

The variable of firm age is based on the registered start year of each firm,

provided also by the Structural Business Statistics. This data tell us firms' start years registered at the tax agency. We set age one to the registered start year. One limitation is that the register only goes back to 1972. Therefore, all firms that started in 1972 or before are assigned an age of 26 in 1997.

The number of employees in the dataset is the annual average converted into full-time equivalents. Value added is deflated with a price index available from the price division of Statistics Sweden. This price index is only defined at the 2-digit level for some industries, while it is defined at the 3-digit or 4-digit level for other industries.

4 Descriptive statistics

Our purpose in this study is to investigate which kinds of firms, in terms of size and age, generate the most employment and productivity growth during the sample period 1997–2006. We will first present descriptive statistics of our dataset.

4.1 Employment

Letting G_n denote a group of firms classified by their size, we compare $\sum_i (L_{i,06} - L_{i,97})$ (for $i \in G_n$) between different G_n. Firstly, the sum of the employment change in the firms that existed in 1997 is expressed by

$$\sum_{i \in C} (L_{i,06} - L_{i,97}) + \sum_{i \in X} (0 - L_{i,97}) \quad \text{(for } i \in G_n\text{)}, \tag{12}$$

where C denotes "continuing firms," i.e., firms that existed in 1997 and survived at least to 2006, and X stands for "exiting firms," i.e., firms that existed in 1997 but went out of business before 2006.[4] One problem in the comparison between different size groups is that the volume of the contribution of each G_n depends on the size of each group. To obtain a measure of the contribution relative to the group size, we divide the expression Eq.(12) by the initial group size, $\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}$:

$$\frac{\sum_{i \in C} (L_{i,06} - L_{i,97})}{\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}} + \frac{\sum_{i \in X} (0 - L_{i,97})}{\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}} = g_n^C + g_n^X \quad \text{(for } i \in G_n\text{)}. \tag{13}$$

---

[4] We follow the terminology used by, for example, Bartelsman and Doms (2000), Foster et al.(2001, 2002), Bartelsman et al.(2004), and Baldwin and Gu (2006).

The term λ_n expresses the unconditional employment growth rate for G_n, which consists of the employment growth in the continuing firms (g_n^C) corresponding to the first term in Eq.(13) and employment loss due to firm exits (g_n^X) corresponding to the second term in Eq.(13). As already discussed, the previous studies focus only on the former part, i.e., the conditional employment growth rate.[5] In contrast, we are interested in analyzing both parts to reveal the net employment change. By investigating the two parts separately, we are also able to tell whether most of the employment change occurred in the continuing firms or was driven by firms that exited the market.

Table 1 reports the numbers of firms and the sums of employees in 1997 by eight size groups. It shows that most firms in the mining and manufacturing industries were small firms with less than ten employees, while nearly half of the employees worked at large firms with more than 500 employees. This unambiguously tells us the importance of large firms in employment.

[Table 1: Firms and employees in 1997]

Table 2 presents the number of firms and the sum of the employees in the firms that survived until 2006. In total, 10,961 firms out of 17,566 were still in business in 2006. We observe a clear trend in the change of $L_{i,t}$, i.e., the number of employees increased in small firms from 1997 to 2006 whereas it decreased in large firms. The employment growth rate g_n^C (see Eq.(13)) clearly shows a negative relationship with firm size.

[Table 2: Employment change in the firms that survived until 2006]

Table 3 shows the number of firms and the sum of the employees in the firms that exited the market before 2006. A total of 6,605 firms out of 17,566 went out of business. Looking at the numbers of lost jobs, we find that more jobs disappeared in larger firms. However, because large firms had more shares of employment initially, the rate of the employment change g_n^X (see Eq.(13)) tends to be somewhat lower for large firms than for small firms.

[Table 3: Employment change in the firms that exited the market before 2006]

---

[5]To be exact, the previous studies analyze the growth rate in the form of $\sum_{i \in C} \left(L_{i,T} - L_{i,0} \right) / \sum_{i \in C} L_{i,0}$, i.e., the denominator only covers the continuing firms.

The total employment growth rate, which is the sum of g_n^C and g_n^X, is presented in Table 4. We observe that it has a clear negative relationship with firm size, implying that employment is more likely to disappear in larger firms. For small firms, employment losses due to firm exits are relatively larger than for large firms. However, as most new jobs are created in small continuing firms, the unconditional growth rate is less negative than for large firms.

[Table 4: Sum of the employment growth rates]

4.2   Labor productivity

In a similar manner, we analyze the productivity contributions of different size groups. Based on Eq.(8), the aggregate change in labor productivity can be expressed as the sum of the contributions from continuing and exiting firms:

$$\sum_{i \in C} \left(LP_{i,06} \cdot s_{i,06} - LP_{i,97} \cdot s_{i,97} \right) + \sum_{i \in X} \left(0 - LP_{i,97} \cdot s_{i,97} \right) \qquad \text{(for } i \in G_n\text{). (14)}$$

To disentangle the relative productivity contribution from group size, we divide it by the initial group size $\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}$:

$$\frac{\sum_{i \in C} \left(LP_{i,06} \cdot s_{i,06} - LP_{i,97} \cdot s_{i,97} \right)}{\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}} + \frac{\sum_{i \in X} \left(0 - LP_{i,97} \cdot s_{i,97} \right)}{\sum_{i \in C} L_{i,97} + \sum_{i \in X} L_{i,97}} = \pi_n^C + \pi_n^X \qquad \text{(for } i \in G_n\text{). (15)}$$

This term expresses the labor productivity contribution per employee of the firms that existed in 1997. It consists of the contributions of the continuing firms π_n^C corresponding to the first term in Eq.(15), and the contributions of the exiting firms π_n^X corresponding to the second terms in Eq.(15).

Table 5 reports the average labor productivity $(\overline{LP}_{i,t})$,[6] the sum of the productivity contributions $(\sum_i LP_{i,t} \cdot s_{i,t})$ of the continuing firms in 1997 and 2006 by size group,[7] as well as the contributions to productivity growth per employee (π_n^C).[8] We observe a clear positive relationship between individual labor productivity $(\overline{LP}_{i,t})$ and firm size. We also find that groups of larger firms have larger figures for the sum of productivity contributions $(\sum_i LP_{i,t} \cdot s_{i,t})$. However, this positive relationship is partly due to larger labor shares of the groups. Dividing the figures by the initial firm group size (in terms

---

[6] The unit is [10^6 Swedish crowns / person].
[7] The unit is [10^6 Swedish crowns / person].
[8] The unit is [Swedish crowns / person^2].

of employees), we obtain π_n^C. In these figures, we observe that the largest per-employee contribution to aggregate productivity growth is found in the group of the largest firms. However, the contributions of small firms are non-negligible, since the second largest contribution to productivity growth is found in the group of the smallest firms. It seems as if there is a U-shaped relationship between growth contribution π_n^C and firm size, though the trend is ambiguous.

[Table 5: Contributions of the firms that existed during the whole period 1997-2006 to labor productivity growth]

Table 6 presents the same information for the firms that exited the market before 2006. The contributions to aggregate productivity growth per employee (π_n^X) suggests that smaller firms have a less negative impact on aggregate productivity. There is also a U-shaped relationship with firm size.

[Table 6: Contributions of the firms that exited the market before 2006 to labor productivity growth]

The total contribution – the sum of π_n^C and π_n^X – subsequently shows in Table 7 that small firms with less than five employees and large firms with more than 500 employees have the most contributions to labor productivity growth per employee, while medium-sized firms ($20 \leq L_{i,97} \leq 499$) contribute marginally.

[Table 7: Sum of the per-employee contributions to labor productivity growth]

5   Econometric analysis

The analysis so far has been based on an arbitrary classification by firm size. Besides, because firms of different ages are mixed in the size groups, more accurate analysis would require a parametric approach.

5.1   Employment

We assume that subsequent employment changes depend on firm size and age at the initial point, i.e., in 1997. The expected level of employment in 2006 conditional on survival is expressed by

$$E\left[L_{i,06}|L_{i,06} > 0\right] = f\left(L_{i,97}, a_{i,97}\right). \tag{16}$$

The unconditional expected level of employment is a product of the conditional expectation and the survival probability:

$$
\begin{aligned}
E\left[L_{i,06}\right] &= E\left[L_{i,06}|L_{i,06} > 0\right] \cdot \Pr\left[L_{i,06} > 0\right] \\
&= f\left(L_{i,97}, a_{i,97}\right) \cdot \Phi\left[h\left(L_{i,97}, a_{i,97}\right)\right], \quad (17)
\end{aligned}
$$

where $f\left(L_{i,97}, a_{i,97}\right) = E\left[L_{i,06}|L_{i,06} > 0\right]$ and $\Phi\left[h\left(L_{i,97}, a_{i,97}\right)\right] = \Pr\left[L_{i,06} > 0\right]$.

Based on these equations, the conditional growth rate of employment g_i^{con} is expressed by

$$
g_i^{con} = \frac{E\left[L_{i,06}|L_{i,06} > 0\right] - L_{i,97}}{L_{i,97}} = \frac{f\left(L_{i,97}, a_{i,97}\right) - L_{i,97}}{L_{i,97}}. \quad (18)
$$

and the unconditional growth rate of employment g_i^{uncon} is expressed by

$$
g_i^{uncon} = \frac{E\left[L_{i,06}\right] - L_{i,97}}{L_{i,97}} = \frac{f\left(L_{i,97}, a_{i,97}\right) \cdot \Phi\left[h\left(L_{i,97}, a_{i,97}\right)\right] - L_{i,97}}{L_{i,97}}, \quad (19)
$$

We parametrize $h\left(\cdot\right)$ and $f\left(\cdot\right)$ as follows:

$$
\begin{aligned}
h\left(L_{i,97}, a_{i,97}\right) &= \beta_0 + \beta_L \ln L_{i,97} + \beta_a a_{i,97} \\
&\quad + \beta_{LL} \left(\ln L_{i,97}\right)^2 + \beta_{aa} \left(a_{i,97}\right)^2 + \beta_{La} \ln L_{i,97} a_{i,97} \\
&\quad + d_j + u_i \quad (20)
\end{aligned}
$$

$$
\begin{aligned}
\ln f\left(L_{i,97}, a_{i,97}\right) &= \gamma_0 + \gamma_L \ln L_{i,97} + \gamma_a a_{i,97} \\
&\quad + \gamma_{LL} \left(\ln L_{i,97}\right)^2 + \gamma_{aa} \left(a_{i,97}\right)^2 + \gamma_{La} \ln L_{i,97} a_{i,97} \\
&\quad + d'_j + v_i \quad (21)
\end{aligned}
$$

where d_j and d'_j are sector dummies for sector j at the 2-digit level, and u_i and v_i are residuals with $iid.\left(0, \sigma_u\right)$ and $iid.\left(0, \sigma_v\right)$, respectively. The benchmark is "furniture, recycling and others" (code 36–37 according to the NACE 2-digit classification). Note that $E\left[L_{i,06}|L_{i,06} > 0\right]$ $\left(= f\left(L_{i,97}, a_{i,97}\right)\right)$ is given by

$$
E\left[L_{i,06}|L_{i,06} > 0\right] = \exp\left(E\left[\ln f\left(L_{i,97}, a_{i,97}\right)\right] + \frac{\sigma_v^2}{2}\right). \quad (22)
$$

5.1.1 Estimation method

For estimation of Eq.(17), we use a lognormal hurdle model, i.e., we estimate the equation by separately running a regression of the equation

$$\Pr\left[L_{i,06} > 0\right] = \Phi\left[h\left(L_{i,97}, a_{i,97}\right) + \epsilon_1\right] \tag{23}$$

using the probit model and a regression of the equation

$$\ln L_{i,06} = f\left(L_{i,97}, a_{i,97}\right) + \epsilon_2 \qquad \text{if } L_{i,06} > 0 \tag{24}$$

using OLS. A separate estimation of the two equations provides us with a consistent estimate of $E\left[L_{i,06}\right]$ on condition that ϵ_1 and ϵ_2 are independent. However, when $L_{i,06}$ is positively related with firm survival so that ϵ_1 and ϵ_2 are positively correlated, the problem of sample selection biases arises. As already mentioned, previous studies have solved the sample selection problem using the Heckman model. The Heckman model implies that we estimate

$$\Pr\left[y_i > 0 | x_i, z_i\right] = \Phi\left[x_i \delta_1 + z_i \delta_2 + \epsilon_{1,i}\right] \tag{25}$$

in the first stage, and

$$y_i = x_i \beta + \mu \lambda\left(x_i \delta_1 + z_i \delta_2\right) + \epsilon_{2,i} \qquad \text{if } y_i > 0 \tag{26}$$

in the second stage, where x_i is a vector of the explanatory variables that are common in both the survival function and the objective function, z_i is variables with exclusion restrictions, and $\lambda\left(\cdot\right) \equiv \phi\left(\cdot\right) / \Phi\left(\cdot\right)$ is the inverse Mills ratio.

In our lognormal hurdle model, we instead estimate

$$\Pr\left[y_i > 0 | x_i\right] = \Phi\left[x_i \tilde{\delta}_1 + \tilde{\epsilon}_{1,i}\right] \tag{27}$$

and

$$y_i = x_i \tilde{\beta} + \tilde{\epsilon}_{2,i} \qquad \text{if } y_i > 0 \tag{28}$$

separately. The reason for our choice of the hurdle model is that we do not have appropriate exclusion restrictions z in the survival function. The hurdle model does not, on the other hand, take into account potential correlation between $\tilde{\epsilon}_{1,i}$ and $\tilde{\epsilon}_{2,i}$. However, in the case where there are no exclusion restrictions, the

hurdle model approximates the Heckman model. Without exclusion restrictions, Eq.(26) is written as

$$y_i = x_i\beta + \mu\lambda(x_i\delta_1) + \epsilon_{2,i} \quad \text{if } y_i > 0. \tag{29}$$

Because the inverse Mills ratio $\lambda(\cdot)$ is monotonously decreasing, we linearly approximate it as $\lambda(x_i\delta_1) \simeq x_i\theta$. Then the right-hand side of Eq.(29) is rewritten as $x_i(\beta + \mu\theta) + \epsilon_{2,i}$. Consequently, the term $\tilde{\beta}$ in the objective function Eq.(28) can be interpreted as reflecting the direct effect β and the effect operating through selection combined.

5.1.2 Result

The estimated values of the survival rate Eq.(23) for different age and firm size are presented in Table 8. They are estimated using the maximum likelihood method. We observe that, for a given size, older firms are more likely to survive, yet the marginal effects become smaller and even negative at higher age. For a given age, the survival rate is highest for medium-sized firms. Note that there are only few observations of very large firms, especially when it comes to newly started firms.

[Table 8: Survival rate for the period 1997-2006]

The conditional growth rate g_i^{con} is presented in Table 9. It is striking that smaller and younger firms grow the most in employment. For a given size, the rate is negatively related with firm age. For a given age, it is negatively related with firm size except for very old firms.

[Table 9: Growth rate of employment, conditional on survival]

The unconditional growth rate g_i^{uncon} is presented in Table 10. Although smaller and younger firms have lower survival rates, their conditional growth rates are much higher than those of other firms. As a result of these two opposite effects, we find that the unconditional growth rate is highest for the small firms $(L_{i,97} < 5)$ at age around five. It can be interpreted that unsucessful firms have already been eliminated in the high mortality phase immediately after entry. Small firms with other ages still have relatively higher growth rates than larger firms. For larger firms, older firms tend to have higher growth rates.

[Table 10: Growth rate of employment, unconditional on survival]

5.2 Labor productivity

A similar econometric analysis is implemented for labor productivity. We assume that labor productivity depends on firm size and age at the initial point. The conditional expected valued of $LP_{i,06}$ is expressed by

$$E\left[LP_{i,06}|L_{i,06} > 0\right] = \psi\left(L_{i,97}, a_{i,97}\right). \tag{30}$$

The unconditional expected value of $LP_{i,06}$ is expressed by

$$
\begin{aligned}
E\left[LP_{i,06}\right] &= E\left[LP_{i,06}|L_{i,06} > 0\right] \cdot \Pr\left[L_{i,06} > 0\right] \\
&= \psi\left(L_{i,97}, a_{i,97}\right) \cdot \Phi\left[h\left(L_{i,97}, a_{i,97}\right)\right]. \tag{31}
\end{aligned}
$$

Besides, since we are interested in labor productivity *growth* for a firm of a given size and age, we also estimate the average level of initial labor productivity $LP_{i,97}$ conditional on $L_{i,97}$ and $a_{i,97}$:

$$E\left[LP_{i,97}\right] = \psi'\left(L_{i,97}, a_{i,97}\right). \tag{32}$$

Based on these equations, the contribution of a firm to productivity growth per employee conditional on firm survival is derived as

$$\pi^{con} = \left(\frac{E\left[LP_{i,06}|L_{i,06} > 0\right] \cdot E\left[L_{i,06}\right]}{\sum_k L_{k,06}} - \frac{LP_{i,97} \cdot L_{i,97}}{\sum_k L_{k,97}}\right) / L_{i,97}, \tag{33}$$

and the contribution of a firm to productivity growth per employee unconditional on firm survival is

$$\pi^{uncon} = \left(\frac{E\left[LP_{i,06}\right] \cdot E\left[L_{i,06}\right]}{\sum_k L_{k,06}} - \frac{LP_{i,97} \cdot L_{i,97}}{\sum_k L_{k,97}}\right) / L_{i,97}. \tag{34}$$

We parametrize $\psi\left(\cdot\right)$ and $\psi'\left(\cdot\right)$ as

$$
\begin{aligned}
\ln \psi\left(L_{i,97}, a_{i,97}\right) &= \delta_0 + \delta_L \ln L_{i,97} + \delta_a a_{i,97} \\
&\quad + \delta_{LL}\left(\ln L_{i,97}\right)^2 + \delta_{aa}\left(a_{i,97}\right)^2 + \delta_{La} \ln L_{i,97} a_{i,97} \\
&\quad + D_j + e_i \tag{35}
\end{aligned}
$$

and

$$
\begin{aligned}
\ln \psi' \left(L_{i,97}, a_{i,97} \right) \quad = \quad & \delta_0' + \delta_L' \ln L_{i,97} + \delta_a' a_{i,97} \\
& + \delta_{LL}' \left(\ln L_{i,97} \right)^2 + \delta_{aa}' \left(a_{i,97} \right)^2 + \delta_{La}' \ln L_{i,97} a_{i,97} \\
& + D_j' + e_i' \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (36)
\end{aligned}
$$

respectively, where D_j and D_j' are sector dummies for sector j at the 2-digit level, and e_i and e_i' are residuals with $iid.\,(0, \sigma_e)$ and $iid.\,(0, \sigma_{e'})$, respectively. The benchmark is again "furniture, recycling and others" (code 36–37 according to the NACE 2-digit classification). Note that $E\left[LP_{i,06}|L_{i,06} > 0\right] \left(= \psi\left(L_{i,97}, a_{i,97}\right)\right)$ and $E\left[LP_{i,97}\right] \left(= \psi'\left(L_{i,97}, a_{i,97}\right)\right)$ are given by

$$
E\left[LP_{i,06}|L_{i,06} > 0\right] = \exp\left(E\left[\ln \psi\left(L_{i,97}, a_{i,97}\right)\right] + \frac{\sigma_e^2}{2} \right) \quad\quad (37)
$$

and

$$
E\left[LP_{i,97}\right] = \exp\left(E\left[\ln \psi'\left(L_{i,97}, a_{i,97}\right)\right] + \frac{\sigma_{e'}^2}{2} \right), \quad\quad (38)
$$

respectively.

5.2.1 Result

We apply the hurdle model in estimating the equations Eq.(31). The conditional productivity contribution σ^{con} is presented in Table 11. We observe that young and small firms contribute the most per employee to aggregate productivity growth. The contributions of large and old firms are also relatively large, yet are only two-thirds of the contributions of small and young firms.[9]

[Table 11: Per-employee contribution to productivity growth, conditional on survival]

Table 12 presents the unconditional productivity contribution σ^{uncon}. The advantage of small and young firms remains even after taking into account their high mortality.

[Table 12: Per-employee contribution to productivity growth, unconditional on survival]

---

[9] The estimated $E\left[LP_{i,06}|L_{i,06} > 0\right]$ and $E\left[LP_{i,97}\right]$ show that labor productivity in both 1997 and 2006 is positively correlated with firm size and age (not presented in this paper).

6 Conclusions

Firms in developed industrial countries face the pressures of globalization. Confronted with cost competition and an increased degree of uncertainty, scale economy becomes less important while knowledge-based economic activities play significant roles in continued economic growth. Several previous studies report that entrepreneurial firms show disproportionately large contributions to employment creation and productivity growth. However, their net effects on the aggregate economy have been unclear due to the high mortality of these young and small firms.

We investigate the importance of entrepreneurial firms in employment generation and productivity growth, unconditional on firm survival. Our study is based on register-based panel data for the Swedish mining and manufacturing industries for the period 1997–2006, with 17,566 legally independent firms with at least one employee.

As a measure of employment contribution, we adapt the growth rate of employment. We found that, although young and small firms are less likely to survive, their employment growth rates conditional on survival are on average higher than those of old and large firms, which is consistent with the previous studies reviewed in this paper. As a result of these opposite effects, the unconditional growth rate of employment is highest for small firms with less than five employees and an age of around five years.

As a measure of productivity contribution, we propose using an individual firm's contribution to the aggregate productivity change per employee. Our choice is motivated because this measure is free from firm size (in terms of number of employees) and thus implies advantages when comparing firms of different sizes. We observe that entrepreneurial firms contribute the most per employee to aggregate productivity growth. Admittedly, also old, large firms exhibit relatively large per-employee contributions to aggregate productivity. However, the magnitude is much less compared with entrepreneurial firms.

Our results can for example be interpreted as follows. Let us assume that a company is failing and that this may result in 100 lost jobs. with potentially 100 job losses. The government faces the choice of either supporting the company and saving the 100 jobs, or letting the company go bankrupt and instead help new firms emerge and in so doing create 100 job opportunities. The question is: what is the relative performance of these two different policies in the long run? More specifically, what are the expected contributions of 50 small firms

with two employees each and of one large firm with 100 employees to the aggregate economy in terms of employment and productivity in ten years? Our study shows that the latter policy may yield better results, in terms of both job opportunities and productivity growth.

References

[1] Acs, Zoltan J. and Audretsch, David B. (1988), "Innovation in Large and Small Firms: An Empirical Analysis," *American Economic Review*, 78(4), 678-690.

[2] Acs, Zoltán J.; Audretsch, David B.; Braunerhjelm, Pontus; and Carlsson, Bo (2005), "The Missing Link: The Knowledge Filter and Entrepreneurship in Endogenous Growth", CEPR Discussion Paper 4783, Centre for Economic Policy Research.

[3] Almus, Matthias and Nerlinger, Eric A. (2000), "Testing "Gibrat's Law" for Young Firms – Empirical Results for West Germany," *Small Business Economics*, 15(1), 1-12.

[4] Audretsch, David B. (2005), "The Emergence of Entrepreneurship Policy," *Local Heroes in the Global Village, International Studies in Entrepreneurship*, 1(7), 21-43.

[5] Audretsch, David B. and Mahmood, Talat (1994), "Firm Selection and Industry Evolution: The Post-entry Performance of New Firms," *Journal of Evolutionary Economics*, 4(3), 243-60.

[6] Audretsch, David B.; Klomp, Luuk; and Thurik, A. Roy (1999a). "Innovation, Industry Evolution, and Employment", In D. Audretsch and R. Thurik (eds.), *Nordic Labour Market Research on Register Data*, 230-52, Cambridge University Press.

[7] Audretsch, David B.; Santarelli, Enrico; and Vivarelli, Marco (1999b), "Start-up size and industrial dynamics: some evidence from Italian manufacturing," *International Journal of Industrial Organization*, 17(7), 965-83.

[8] Audretsch, David B. and Thurik, A. Roy (2000), "Capitalism and Democracy in the 21st Century: From the Managed to the Entrepreneurial Economy," *Journal of Evolutionary Economics*, 10(1), 17-34.

[9] Audretsch, David B. and Thurik, A. Roy (2001), "Linking Entrepreneurship to Growth," OECD STI Working Paper 2081/2.

[10] Baldwin, John R. and Picot, Garnett (1995), "Employment generation by small producers in the Canadian manufacturing sector," *Small Business Economics*, 7(4), 317-31.

[11] Baldwin, John R. and Gu, Wulong (2006), "Plant turnover and productivity growth in Canadian manufacturing." *Industrial and Corporate Change*, 15(3), 417-65.

[12] Balk, Bert M. and Hoogenboom-Spijker, Ellen (2003), "The measurement and decomposition of productivity change: exercises on the Netherlands' manufacturing industry," Statistic Netherlands, discussion paper 03001.

[13] Barnes, Matthew and Haskel, Jonathan (2000), "Productivity, competition and downsizing," Queen Mary, University of London, HM Treasury Growth seminar, London

[14] Bartelsman, Eric; Haltiwanger, John; and Scarpetta, Stefano (2004), "Microeconomic evidence of creative destruction in industrial and developing countries," Policy Research Working Paper Series 3464, The World Bank.

[15] Bartelsman, Eric J. and Doms. Mark E. (2000), "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 38(3), 569-94.

[16] Baumol, William J. (2002), *The Free-market Innovation Machine - Analysing the Growth Miracle of Capitalism of Last Year*, Princeton University Press, Princeton.

[17] Baumol, William J. (2005) "Small Firms: Why Market-Driven Innovation Can't Get Along without Them", U.S. Small Business Administration.

[18] Biggs, Tyler (2002), 'Is small beautiful and worthy of subsidy?' International Finance Corporation, Washington.

[19] Brock, William A. and Evans, David S. (1989), "Small business economics," *Small Business Economics*, 1(1), 7-20.

[20] Broersma, Lourens and Gautier, Pieter (1997), "Job creation and job destruction by small firms: An empirical investigation for the Dutch manufacturing sector," *Small Business Economics*, 9(3), 211-24.

[21] Burgess, Simon; Lane, Julia; and Stevens, David (2000), "Job flows, worker flows, and churning," *Journal of Labor Economics*, 18(3), 473-502.

[22] Callejon, Maria and Segarra, Agusti (1999), "Business dynamics and efficiency in industries and regions: The case of Spain," *Small Business Economics*, 13(4), 253-71.

[23] Calvo, José L. (2006), "Testing Gibrat's law for small, young, and innovating firms," *Small Business Economics*, 26(2), 117-23.

[24] Carlsson, Bo (1992), "The Rise of Small Business: Causes and Consequences," in Adams, William J. (ed.), *Singular Europe, Economy and Policy of the European Community after 1992*, Ann Arbor, University of Michigan Press, 145-69.

[25] Carree, Martin A. (2002), "Industrial restructuring and economic growth", *Small Business Economics*, 18(4), 243-55.

[26] Carree, Martin A. and Klomp, Luuk (1996), "Small Business and Job Creation: A Comment," *Small Business Economics*, 8(4), 317-22.

[27] Carree, Martin A. and Thurik, A. Roy (2006), "The lag structure of the impact of business ownership on economic performance in OECD countries," *Small Business Economics*, 30(1), 101-10.

[28] Castany, Laia; López-Bazo, Enrique; and Moreno, Rosina (2005), "Differences in total factor productivity across firm size. A distributional analysis", University of Barcelona Working Paper.

[29] Celikkol, Pinar (2003), "Productivity patterns in the U.S. food and kindred products industries: a plant level analysis, 1972–1995," The Pennsylvania State University.

[30] Davis, Steven J. and Haltiwanger, John C. (1992), "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 107(3), 819-63.

[31] Davis, Steven J.; Haltiwanger, John C.; and Schuh, Scott (1996), "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts," *Small Business Economics*, 8(4), 297-315.

[32] de Kok, Jan; Brouwer, Peter; and Fris, Pieter (2005), "Can firm age account for productivity differences?,' Scales Research Reports N200421, EIM Business and Policy Research.

[33] Disney, Richard; Haskal, Jonathan E.; and Heden, Ylva (2003), "Restructuring and productivity growth in UK manufacturing," *The Economic Journal*, 113(489), 666-94.

[34] Dunne, Paul and Hughes, Alan (1994), "Age, Size, Growth and Survival: UK Companies in the 1980s," *Journal of Industrial Economics*, 42(2), 115-40.

[35] Dunne, Timothy; Roberts, Mark J. and Samuelson, Larry (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, 19(4), 495-515.

[36] Ericson, Richard and Pakes, Ariel (1992), "An Alternative Theory of Firm and Industry Dynamics," Cowles Foundation Discussion Papers 1041, Cowles Foundation for Research in Economics, Yale University.

[37] Evans, David S. (1987a), "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries," *Journal of Industrial Economics*, 35(4), 567-81.

[38] Evans, David S. (1987b), "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy*, 95(4), 657-74.

[39] Foster, Lucia; Haltiwanger, John C.; and Krizan, C. J. (2001), "Aggregate Productivity Growth. Lessons from Microeconomic Evidence", in Hulten, C. R.; Dean, E. R.; and Harper, M. J. (eds.), *New Developments in Productivity Analysis*, National Bureau of Economic Research, 303-72.

[40] Foster, Lucia; Haltiwanger, John C.; and Krizan, C. J. (2002), "The Link Between Aggregate and Micro Productivity Growth: Evidence from Retail Trade," NBER Working Papers 9120.

[41] Fritsch, Michael (2008), "How does new business formation affect regional development?," *Small Business Economics*, 30(1), 1-14.

[42] Fölster, Stefan (2000), "Do Entrepreneurs Create Jobs?," *Small Business Economics*, 14(2), 137-48.

[43] Geroski, Paul A. (1995), *Market Structure, Corporate Performance and Innovative Activity*, Clarendon Press, Oxford.

[44] Gibrat, Robert (1931), Les Ine´galite´s Economiques. Paris, France: Librairie du Recueil Sirey.

[45] Goddard, John; Wilson, John; and Blandon, Peter (2002), "Panel tests of Gibrat's Law for Japanese manufacturing," *International Journal of Industrial Organization*, 20(3), 415-33.

[46] Hall, Bronwyn H. (1987), "The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector," *Journal of Industrial Economics*, 35(4), 583-606.

[47] Haltiwanger, John C.; Jarmin, Ron S. and Miranda, Javier (2010), "Who Creates Jobs? Small vs. Large vs. Young", NBER Working Papers 16300.

[48] Hart, Peter E. and Oulton, Nicholas (1996), Growth and size of firms, *The Economic Journal*, 106(438), 1242–52.

[49] Heckman, James J., (1979), "Sample selection bias as a specification error," *Econometrica*, 47(1), 153-61.

[50] Heshmati, Almas (2001), "On the Growth of Micro and Small Firms: Evidence from Sweden," *Small Business Economics*, 17(3), 213-28.

[51] Huergo, Elena and Jaumandreu, Jordi (2004), "How does probability of innovation change with firm age?," *Small Business Economics*, 22(3-4), 193-207.

[52] Jensen, J. Bradford; McGuckin, Robert H.; and Stiroh, Kevin J. (2001), "The impact of vintage and survival on productivity: Evidence from cohorts of U.S. manufacturing plants," *The Review of Economics and Statistics*, 83(2), 323-32.

[53] Johansson, Dan (2005), "The turnover of firms and industry growth," *Small Business Economics*, 24(5), 487-95.

[54] Jovanovic, Boyan (1982), "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649-70.

[55] Konings, Jozef (1995), "Gross job flows and the evolution of size in U.K. establishments," *Small Business Economics*, 7(3), 213-20.

[56] Lever, M. H. C. (1996), "Firm size and employment determination in Dutch manufacturing industries," *Small Business Economics*, 8(5), 389-96.

[57] Lindvall, Johannes and Rueda, David (2010), "The Insider-Outsider Dilemma in Sweden," Dualisation Workshop, Oxford University, January 14-16, 2010.

[58] Loveman, Gary and Sengenberger, Werner (1991), "The re-emergence of small-scale production: An international comparison," *Small Business Economics*, 3(1), 1-31.

[59] Lundström, Anders and Stevenson, Lois A. (2005), *Entrepreneurship Policy. Theory and Practice, International Studies in Entrepreneurship*, Springer, Boston.

[60] Mansfield, Edwin (1962), "Entry, Gibrat's Law, Innovation, and the Growth of Firms," *The American Economic Review*, 52(5), 1023-51.

[61] Mata, José (1994), "Firm growth during infancy," *Small Business Economics*, 6(1), 27-39.

[62] Nurmi, Satu (2004), "Plant Size, Age and Growth in Finnish Manufacturing," *Finnish Economic Papers*, 17(1), 3-17.

[63] Nguyen, Sang V. and Lee, Seong-Hoon (2002), "Returns to scale in small and large U.S. manufacturing establishments: Further evidence," *Small Business Economics*, 19(1), 41–50.

[64] Oliveira, Blandina and Fortunato, Adelino (2006), "Firm growth and liquidity constraints: A dynamic analysis," *Small Business Economics*, 27(2-3), 139-56.

[65] Picot, Garnett and Dupuy, Richard (1998), "Job creation by company size class: The magnitude, concentration and persistence of job gains and losses in Canada," *Small Business Economics*, 10(2), 117-39.

[66] Power, Laura (1998), "The missing link: technology, investment and productivity," *Review of Economics and Statistics*, 80(2), 300-13.

[67] Robbins, D. Keith; Pantuosco, Louis J.; Parker, Darrell, F.; and Fuller, Barbara, K. (2000), "An empirical assessment of the contribution of small business employment to U.S. state economic performance," *Small Business Economics*, 15(4), 293-302.

[68] Rodríguez, Alicia Correa; Molina, Miguel Acosta; Ana L. González Pérez; and Urbano Medina Hernández (2003), "Size, age and activity sector on the growth of the small and medium firm size," *Small Business Economics*, 21(3), 289-307.

[69] Shaffer, Sherrill (2006), "Establishment size and local employment growth," Small Business Economics, 26(5), 439-54.

[70] Storey, David (2005), "Entrepreneurship, Small and Medium Sized Enterprises and Public Policies," *Handbook of Entrepreneurship Research, International Handbook Series on Entrepreneurship*, 1(7), 473-511.

[71] Sutton, John (1997), "Gibrat's legacy," *Journal of Economic Literature*, 35(1), 40-59.

[72] van Praag; Mirjam, C.; and Versloot, Peter H. (2007), "What is the value of entrepreneurship? A review of recent research," *Small Business Economics*, 29(4), 351-382.

[73] Taymaz, Erol (2005), "Are Small Firms Really Less Productive?," *Small Business Economics*, 25(5), 429-45.

[74] Wennekers, Sander and Thurik, A. Roy (1999), "Linking Entrepreneurship and Economic Growth," *Small Business Economics*, 13(1), 27-55.

Size groups based on firm size in 1997	All firms in 1997	
	[1] = [3]+[7]	[2] = [4]+[8]
	# of firms	Sum ($L_{i,97}$)
$1 \leq L < 5$	7 537	17 264
$5 \leq L < 10$	3 839	25 356
$10 \leq L < 20$	2 609	35 328
$20 \leq L < 50$	1 949	58 553
$50 \leq L < 100$	782	54 428
$100 \leq L < 200$	409	56 539
$200 \leq L < 500$	259	77 925
$L \geq 500$	182	270 557
Total	17 566	595 950

Table 1: Firms and employees in 1997

Size groups based on firm size in 1997	Firms that survived until 2006 (Continuing firms)			
	[3]	[4]	[5]	[6] = ([5]–[4]) / [2]
	# of firms	Sum ($L_{i,97}$)	Sum ($L_{i,06}$)	Growth rate: g^{C}_{n}
$1 \leq L < 5$	4 463	10 671	15 123	0.258
$5 \leq L < 10$	2 535	16 764	20 421	0.144
$10 \leq L < 20$	1 672	22 585	27 808	0.148
$20 \leq L < 50$	1 212	36 757	43 784	0.120
$50 \leq L < 100$	518	36 297	43 186	0.127
$100 \leq L < 200$	266	36 901	39 328	0.043
$200 \leq L < 500$	176	52 335	49 656	-0.034
$L \geq 500$	119	184 712	174 371	-0.038
Total	10 961	397 022	413 677	0.028

Table 2: Employment change in the firms that survived until 2006

Size groups based on firm size in 1997	Firms that exited the market before 2006 (Exiting firms)			
	[7]	[8]	[9]	[10] = ([9]–[8]) / [2]
	# of firms	Sum ($L_{i,97}$)	Sum ($L_{i,06}$)	Growth rate: g^{X}_{n}
$1 \leq L < 5$	3 074	6 593	0	-0.382
$5 \leq L < 10$	1 304	8 592	0	-0.339
$10 \leq L < 20$	937	12 743	0	-0.361
$20 \leq L < 50$	737	21 796	0	-0.372
$50 \leq L < 100$	264	18 131	0	-0.333
$100 \leq L < 200$	143	19 638	0	-0.347
$200 \leq L < 500$	83	25 590	0	-0.328
$L \geq 500$	63	85 845	0	-0.317
Total	6 605	198 928	0	-0.334

Table 3: Employment change in the firms that exited the market before 2006

Size groups based on firm size in 1997	[11] = [6]+[10] Total growth rate: $g^{C}_{n} + g^{X}_{n}$
$1 \leq L < 5$	-0.124
$5 \leq L < 10$	-0.195
$10 \leq L < 20$	-0.213
$20 \leq L < 50$	-0.252
$50 \leq L < 100$	-0.207
$100 \leq L < 200$	-0.304
$200 \leq L < 500$	-0.363
$L \geq 500$	-0.356
Total	-0.306

Table 4: Sum of the employment growth rates

| | Firms that survived until 2006 (Continuing firms) | | | | | |
Size groups classified based on firm size in 1997	[12] # of firms	[13] Mean ($LP_{i,97}$)	[14] Sum ($LP_{i,97} \cdot s_{i,97}$)	[15] Mean ($LP_{i,06}$)	[16] Sum ($LP_{i,06} \cdot s_{i,06}$)	[17] = ([16]−[14]) / [2] Contribution to productivity growth per employee: π^C_n
$1 \le L < 5$	4 463	0.366	0.006	0.418	0.012	0.326
$5 \le L < 10$	2 535	0.379	0.011	0.451	0.017	0.261
$10 \le L < 20$	1 672	0.417	0.016	0.524	0.027	0.320
$20 \le L < 50$	1 212	0.432	0.027	0.509	0.042	0.254
$50 \le L < 100$	518	0.456	0.028	0.551	0.043	0.281
$100 \le L < 200$	266	0.497	0.031	0.796	0.045	0.245
$200 \le L < 500$	176	0.533	0.047	0.704	0.066	0.238
$L \ge 500$	119	0.624	0.203	0.809	0.304	0.375
Total	10 961	0.397	0.368	0.476	0.556	0.315

Table 5: Contributions of the firms that existed during the whole period 1997-2006 to labor productivity growth

| | Firms that exited the market before 2006 (Exiting firms) | | | | | |
Size groups classified based on firm size in 1997	[18] # of firms	[19] Mean ($LP_{i,97}$)	[20] Sum ($LP_{i,97} \cdot s_{i,97}$)	[21] Mean ($LP_{i,06}$)	[22] Sum ($LP_{i,06} \cdot s_{i,06}$)	[23] = ([22]−[20]) / [2] Contribution to productivity growth per employee: π^X_n
$1 \le L < 5$	3 074	0.362	0.004	0	0	-0.225
$5 \le L < 10$	1 304	0.389	0.006	0	0	-0.221
$10 \le L < 20$	937	0.396	0.008	0	0	-0.239
$20 \le L < 50$	737	0.411	0.015	0	0	-0.259
$50 \le L < 100$	264	0.444	0.014	0	0	-0.250
$100 \le L < 200$	143	0.467	0.015	0	0	-0.272
$200 \le L < 500$	83	0.472	0.020	0	0	-0.262
$L \ge 500$	63	0.429	0.059	0	0	-0.219
Total	6 605	0.385	0.142	0	0	-0.238

Table 6: Contributions of the firms that exited the market before 2006 to labor productivity growth

Size groups classified based on firm size in 1997	[24] = [17]+[23] $\pi^C_n + \pi^X_n$
$1 \le L < 5$	0.101
$5 \le L < 10$	0.039
$10 \le L < 20$	0.081
$20 \le L < 50$	-0.005
$50 \le L < 100$	0.031
$100 \le L < 200$	-0.027
$200 \le L < 500$	-0.024
$L \ge 500$	0.156
Total	0.077

Table 7: Sum of the per-employee contributions to labor productivity growth

(a) Survival rate

Firm size in 1997:						
L =	age =	1	5	10	20	26
1 ≤ L < 5		0.441	0.502	0.554	0.580	0.546
5 ≤ L < 10		0.480	0.550	0.612	0.658	0.638
10 ≤ L < 20		0.484	0.557	0.624	0.678	0.664
20 ≤ L < 50		0.479	0.557	0.628	0.690	0.682
50 ≤ L < 100		0.462	0.544	0.621	0.695	0.694
100 ≤ L < 200		0.439	0.525	0.608	0.691	0.695
200 ≤ L < 500		0.409	0.498	0.586	0.680	0.689
L ≥ 500		0.358	0.450	0.545	0.654	0.670

Table 8: Survival rate for the period 1997-2006

(b) Conditional growth rate of employment

Firm size in 1997:						
L =	age =	1	5	10	20	26
1 ≤ L < 5		0.873	0.661	0.471	0.270	0.236
5 ≤ L < 10		0.524	0.373	0.240	0.113	0.108
10 ≤ L < 20		0.419	0.287	0.172	0.069	0.076
20 ≤ L < 50		0.336	0.219	0.120	0.039	0.056
50 ≤ L < 100		0.253	0.154	0.071	0.016	0.046
100 ≤ L < 200		0.208	0.120	0.049	0.011	0.052
200 ≤ L < 500		0.177	0.099	0.037	0.017	0.069
L ≥ 500		0.156	0.088	0.039	0.042	0.109

Table 9: Growth rate of employment, conditional on survival

(c) Unconditional growth rate in employment

Firm size in 1997:						
L =	age =	1	5	10	20	26
1 ≤ L < 5		-0.174	-0.167	-0.185	-0.263	-0.325
5 ≤ L < 10		-0.268	-0.245	-0.241	-0.268	-0.293
10 ≤ L < 20		-0.313	-0.283	-0.269	-0.275	-0.286
20 ≤ L < 50		-0.360	-0.321	-0.297	-0.283	-0.280
50 ≤ L < 100		-0.422	-0.373	-0.334	-0.293	-0.274
100 ≤ L < 200		-0.470	-0.412	-0.363	-0.301	-0.269
200 ≤ L < 500		-0.519	-0.453	-0.392	-0.308	-0.264
L ≥ 500		-0.586	-0.510	-0.433	-0.319	-0.257

Table 10: Growth rate of employment, unconditional on survival

(d) Conditional contribution to productivity growth per employee

Firm size in 1997:						
L =	age =	1	5	10	20	26
1 ≤ L < 5		0.685	0.508	0.348	0.191	0.183
5 ≤ L < 10		0.606	0.444	0.299	0.172	0.188
10 ≤ L < 20		0.580	0.424	0.285	0.171	0.199
20 ≤ L < 50		0.559	0.408	0.275	0.175	0.216
50 ≤ L < 100		0.539	0.395	0.269	0.189	0.250
100 ≤ L < 200		0.531	0.392	0.272	0.210	0.288
200 ≤ L < 500		0.530	0.396	0.282	0.239	0.338
L ≥ 500		0.539	0.412	0.307	0.295	0.427

Table 11: Per-employee contribution to productivity growth, conditional on survival

(e) Unconditional contribution to productivity growth per employee

Firm size in 1997:						
L =	age =	1	5	10	20	26
1 ≤ L < 5		0.061	0.022	-0.030	-0.104	-0.119
5 ≤ L < 10		0.030	-0.001	-0.044	-0.093	-0.088
10 ≤ L < 20		0.005	-0.020	-0.057	-0.092	-0.075
20 ≤ L < 50		-0.026	-0.045	-0.074	-0.093	-0.063
50 ≤ L < 100		-0.078	-0.087	-0.103	-0.096	-0.046
100 ≤ L < 200		-0.125	-0.126	-0.131	-0.101	-0.033
200 ≤ L < 500		-0.180	-0.172	-0.165	-0.107	-0.020
L ≥ 500		-0.265	-0.246	-0.220	-0.121	-0.004

Table 12: Per-employee contribution to productivity growth, unconditional on survival

Paper 5

Initial firm size and post-entry growth in size and productivity

Yoshihiro Sato
University of Gothenburg

Abstract

Previous studies show that smaller entrants exhibit higher growth rates in terms of size (number of employees). Using register-based firm-level data for the Swedish mining and manufacturing industries, this study compares the development in size and productivity between a group of firms that started their business in 1998 and a group of firms that had been in business for at least 10 years in 1998. The results show that there is also a similar negative relationship between firms' initial size and post-entry productivity development. The average total factor productivity of the entrants is initially 15 percent lower than that of the incumbents and the difference becomes insignificant after three years. Regarding the growth pattern conditional on initial size, the entrants with one initial employee caught up with the incumbents of similar initial size already in the second year, and gained a lead in the fifth year. It takes three years for the productivity gap between the entrants with 10 initial employees and the incumbents with the same size to disappear. The entrants with more than 20 initial employees never caught up with the entrants of the same size during the nine years analyzed in this study.

1 Introduction

Successful development of productive firms and their international competitiveness are in the spotlight of policy discussions. Since the era of large and capital-intensive industries is over, policymakers are increasingly focusing on the potentials of small and new firms. A smooth and successful reallocation of productive resources from exiting to entering firms is regarded as a key driving force for sustainable economic growth.

Among all Swedish manufacturing and mining firms with at least one employee during the period 1997-2006, newly started firms and exiting firms each

year make up around 6.7 and 5.6 percent, respectively. As a result, 41.8 percent of all the 18,855 firms with at least one employee in 2006 are new firms that entered the market during 1998–2006, while 37.6 percent of all the 17,566 firms that operated in 1997 exited the market sometime from 1997 to 2005 and hence were no longer operating in 2006. Since it is commonly known that the productivity of new firms is on average lower than that of old firms (e.g., Power, 1998; Jensen et al., 2001; and de Kok et al., 2005; see also Chapter 4 in this thesis), this firm turnover may positively affect the development of aggregate productivity as long as newly started firms catch up with incumbent firms in terms of productivity.

The speed of catching up may differ between firms of different initial sizes. Previous studies, although their analyses are not limited to new firms, report that the expected firm size growth rate is negatively related to current size (e.g., Audretsch and Mahmood, 1994; Hart and Oulton, 1996; Audretsch et al., 1999a, 1999b; Heshmati, 2001). This finding suggests that smaller firms grow faster than medium-sized and large firms. However, it is not clear whether the productivity of initially small firms really converges to that of medium-sized and large firms.

Catching up of new firms in terms of productivity also matters for successful reallocation of resources in the economy. Chapter 3 in this thesis analyzes scale elasticity of the production function for the Swedish mining and manufacturing industries during the period 1997–2006. It is found that micro firms (fewer than 10 employee) and small firms (10–50 employees) enjoy increasing returns to scale while larger firms face slightly decreasing returns to scale. This finding implies that firms starting out small have a larger potential for productivity gains as long as they survive and grow in size.

The purpose of this study is to analyze post-entry performance of new firms in terms of size and total factor productivity (TFP) and its relationship with initial firm size. We address entrants' performance in relation to the incumbent firms that belong to the same size class at the start of the sample period.

The question of how long it takes for new firms to catch up with market incumbents is also relevant to the discussion on so-called enterprise zones, or empowerment zones. These zones have been a policy instrument in the U.S. and some European countries in the last decades to stimulate economic development in economically and socially distressed urban areas. The programs typically provide a variety of tax incentives to firms in target areas during a limited period. Such policy includes, for example, tax credits related to creating jobs for

economically disadvantaged workers and capital investments in the designated areas (Kolko and Neumark, 2009; Hanson, 2009; Neumark and Kolko, 2010).

In Sweden, the government has proposed an introduction of similar enterprise zones, or so-called "new start zones" with tax credits for business activities in specified urban regions. The purpose is to tackle social and economic problems in those areas by promoting employment and integrating residents into the labor market (SOU 2012:50).

The introduction of enterprise zones raises the question of how long the policy should be given to firms in the target areas. Many existing enterprise zones provide tax credits and other kinds of support during the first five years (Neumark and Kolko, 2010). In the proposal provided by the investigation committee to the Swedish government, new firms in the target areas will receive a full exemption of payroll tax in the first five years and a 50 percent exemption in the following two years.

Using register-based firm-level data for the Swedish mining and manufacturing industries, we compare the development of size, in terms of number of employees, and productivity between a group of firms that started their business in 1998 (the entrant group) and a group of firms that had been in business for at least 10 years in 1998 (the incumbent group). The results show that the average size of the new entrants grows from 38 to 60 percent of that of the incumbents in eight years. We also find a negative relationship between initial size and post-entry size growth, which is consistent with previous studies.

What is new in the present paper is that a similar negative relationship is found between initial firm size and post-entry productivity development. With all entrants counted, their average TFP is 15 percent lower than the average TFP of the incumbents, and the difference becomes insignificant after three years. However, we find a large difference in growth pattern depending on initial firm size. The entrants with one initial employee caught up with the incumbents of similar initial size already in the second year, and gained a lead in the fifth year. It takes three years for the productivity gap between the entrants with 10 initial employees and the incumbents with the same size to disappear. The entrants with more than 20 initial employees never caught up with the entrants of the same size during the nine years analyzed in this study.

Section 2 discusses relevant previous studies. Descriptions on our dataset is found in Section 3, which is followed by illustrations of basic facts on the pattern of entry and exit and the distribution of firm size in Sweden in Section 4. In Section 5, we focus on two groups, the entrant and the incumbent groups, and

compare the developments of firm size, labor productivity and TFP. Section 6 concludes the paper.

2 Previous studies

The current paper studies the relative positions of new firms in terms of size and productivity compared with incumbent firms in the market. One of the previous studies is Baldwin and Rafiquzzaman (1995), who analyze greenfield entries in the Canadian manufacturing firms 1970–1989. They find that labor productivity and average wages in the first year after entry are 60 and 70 percent of the respective sector averages. After 10 years, the measures reach only 80 percent of the sector averages. However, this is explained by the fact that new and small firms are less capital intensive than large firms. In contrast, the entrants' profitability (i.e., value added minus labor costs) is only 10 percent below the sector average.

Bellone et al. (2006, 2008) analyze French manufacturing firms with more than 20 employees during the period 1990–2002. They find that, conditional on survival, entrant firms on average enjoy continuous productivity gains in terms of TFP during the whole sample period after entry. They also find that the entrants initially have a significant productivity advantage over incumbents. Their TFP measures in the first year are on average 0.6 percent higher than the industry average. It then rises to 1.25 percent at age three before monotonously converging toward the industry average.

A previous study that uses Swedish microdata to analyze post-entry performance is Heshmati (2001). He investigates the growth pattern of both new and old firms with 1–100 employees that operated in the county of Gävleborg during the period 1993–1998. The focus of his study is on firm size in terms of employees, sales, and assets, while productivity is not addressed. Another related study, Persson (2004), analyzes the survival pattern and employment growth of new establishments created in Sweden in 1987 and 1988. Both studies report similar results as previous studies in other economies.

3 Data

The present study is based on firm-level panel data from Statistics Sweden's Structural Business Statistics. The same database is used in Chapters 3 and 4

of this thesis. More details about the database is found in Chapter 3.

In the present study, a firm entry is defined when a firm appears in the database for the first time. In 1997, i.e., the start year of the sample period, it is not possible to distinguish the firms that existed in 1996 from the firms that were established in 1997. Therefore, firm entry is first counted in 1998. Similarly, a firm exit is defined when a firm no longer appears in the following year. In 2006, i.e., the end year of the sample period, it is not possible to distinguish between the firms that still existed in 2007 and the firms that went out of business in 2006. Thus, the last exiting firms are counted in 2005. There is another limitation in the database as well. It is not possible to distinguish between entries resulting from new startups and entries resulting from establishment of subsidiaries by other firms. Neither is it possible to distinguish between exits due to closures or bankruptcy and exits due to acquisition or merger.

Capital costs are defined as user costs, which are calculated as $c_{j,t} = (r_t + \delta_{j,t})/(1 - u_t)$ where r_t denotes the interest rate (defined as the annual average of the official discount rate, i.e., the repo rate), $\delta_{j,t}$ the depreciation rate, and u_t the corporate tax rate (0.28 during the whole sample period). The depreciation rates are deduced from the Swedish national accounts and are both sector specific at the 2-digit level and object specific (buildings and machines).[1]

The registered start year of each firm is also provided by the Structural Business Statistics. This data tell us firms' start years registered at the tax agency.

4 Entry and exit

4.1 Over time

Table 1 presents key statistics on entries and exits of Swedish mining and manufacturing firms. Between 1,000 and 1,400 firms are started every year, making up 5–8 percent of the total number of firms (roughly 18,000). The number of exits per year decreases from 1,400 to 800 during our sample period, which can be explained by the economic boom in 2002–2005. The table shows that 4–8 percent of all firms exit every year.

[Table 1: Entry and exit in the Swedish mining and manufacturing industries]

---

[1] The sector-specific depreciation rate varies between 0.032 and 0.077 for buildings and between 0.123 and 0.437 for machines.

The labor shares of entry firms and exit firms are also presented in Table 1. Labor is defined as the number of employees. The labor share of the entry firms varies from 2 to 5 percent and is much smaller than the entry rate, suggesting that entry firms are generally smaller in the beginning. The labor share of the exit firms is also smaller than the exit rate except for in 1999 and 2001, implying that relatively large firms left the market in the two years.

Table 2 reports the average size of all firms and of entry firms and exit firms by year. It is observed that firms on average become smaller during the sample period. Again, the average entry size is smaller than the average size of all firms, and also tends to decrease over time. The average exit size is larger than that of the entry firms. Again, the table shows that relatively large firms exited in 1999 and 2001. A closer look at the original data reveals that closures of large firms in the automobile and food industries raised the average exit size in 1999. In 2001, it was the industries of electronic machinery and fabricated metal products that contributed to the rise.

[Table 2: Average size (in terms of employees) of all firms, entry firms and exit firms by year]

4.2   By sector

Table 3 presents the entry and exit rates and the average sizes of all firms and of entry firms and exit firms by sector. The entry and exit rates vary between 0.045 and 0.095 and between 0.036 and 0.079, respectively, across sectors. The coefficient of correlation of the two figures is 0.655, which indicates that a sector with a high entry rate is associated with a high exit rate.

[Table 3: Entry rates, exit rates, and average firm size by industry]

As Table 3 shows, the average firm size differs considerably across sectors. This heterogeneity can be explained by the theory of the minimum efficient scale (MES). According to this theory, each well-defined sector has own characteristics of economies of scale and the optimal size (Dunne et al., 1988, 1989; Audretsch, 1995).

4.3 Cohort survival

Table 4 presents the number of survival firms (upper) and the survival rate (lower) subsequent to entry. Of the approximately 1,000–1,400 firms that enter each year, 83–92 percent survive at least the first year; 61-73 percent survive at least four years.

[Table 4: Cohort survival]

5 Post-entry performance of the entrant group in relation to the incumbent group

We now move to post-entry performance in terms of size (number of employees) and TFP. We compare the development of a group of entrants and a group of incumbents from 1998 to 2006. The entrant group includes all firms that entered the market in 1998, i.e., the firms that first appear in our dataset in 1998. The incumbent group includes all firms whose start year registered at the tax agency is 1988 or earlier, i.e., the firms with an age of at least 10 years in 1998.

Table 5 describes basic statistics on the two groups. The number of incumbents is 8,862 in 1998 and reduces to 6,306 in 2006 (implying a survival rate of 71 percent). The entrant group starts with 1,170 firms in 1998 and ends with 589 firms in 2006 (a survival rate of 50 percent).

[Table 5: Comparison of the two groups]

The current study uses TFP as a measure of productivity. The following definition is adapted following previous major studies (Caves et al., 1982; Good et al., 1997; Bellone et al., 2006, 2008; Aw et al., 2001):

$$
\begin{aligned}
\ln TFP_{i,t} \;=\; & \left(\ln Y_{i,t} - \overline{\ln Y_t} \right) + \sum_{\tau=2}^{t} \left(\overline{\ln Y_\tau} - \overline{\ln Y_{\tau-1}} \right) \\
& - \left[
\begin{array}{l}
\sum_{n=1}^{N} \frac{1}{2} \left(S_{i,t}^n + \overline{S_t^n} \right) \left(\ln X_{i,t}^n - \overline{\ln X_t^n} \right) \\
+ \sum_{\tau=2}^{t} \sum_{n=1}^{N} \frac{1}{2} \left(\overline{S_\tau^n} + \overline{S_{\tau-1}^n} \right) \left(\overline{\ln X_\tau^n} - \overline{\ln X_{\tau-1}^n} \right)
\end{array}
\right] ,
\end{aligned}
\tag{1}
$$

where Y denotes value added, X^n denotes n th input (capital stocks K and labor L in the current analysis), and S^n is the cost share of input X^n in the

total cost. The subscript t stands for time (year) and n is an index for different inputs. Upper bars denote sample means. Note that the sample means of the cost shares, $\overline{S_t^n}$, is the arithmetic means, while the sample means of log value added and log input, $\overline{\ln Y_t}$ and $\overline{\ln X_{i,t}^n}$, are the geometric means. Or, we can obtain the same values of $\overline{\ln Y_t}$ and $\overline{\ln X_{i,t}^n}$ if we first compute the arithmetic means of Y_t and $X_{i,t}^n$ and then take log of the arithmetic means.

The first term of Eq.(1) measures the deviation of value added of individual firms from the mean value added in that year. The third term (the one on the second line) sums the deviation of inputs used by individual firms from the mean in that year. The average of the individual cost share and the mean cost share is used as weight. The combination of the first term and (*minus*) the third term measures the cross-sectional distribution of TFP in a year. The second and fourth terms sum the change in the mean TFP across years, by chain-linking the shift of the TFP distribution over time. The whole index $\ln TFP_{i,t}$ indicates a relative location of the TFP of an individual firm i in time t relative to the mean TFP in the reference year.

5.1 Firm size

We first compare the development of average firm size (in terms of the number of employees) between the two groups. We use a simple model with dummy variables for groups, years, and sectors:

$$\ln L_{i,t} = \sum_t \alpha_{0,t} D_{i,t}^{year} + \sum_t \alpha_{1,t} D_{i,t}^{ent} D_{i,t}^{year} + \lambda_i^s + \epsilon_{i,t}, \tag{2}$$

where $L_{i,t}$ is the number of employees, $D_{i,t}^{year}$ is a dummy variable, $D_{i,t}^{ent}$ is a dummy variable for the entrant group, and λ_i^s is the sector dummy at the 3-digit level.

The results are presented in Table 6. The average number of employees of the incumbents is

$$E\left[L_t^{inc}\right] = \exp\left(\alpha_{0,t} + \sigma_\epsilon^2/2\right), \tag{3}$$

and that of the entrants is

$$E\left[L_t^{ent}\right] = \exp\left(\alpha_{0,t} + \alpha_{1,t} + \sigma_\epsilon^2/2\right), \tag{4}$$

where σ_ϵ^2 stands for the mean square of errors. The table shows that, while the incumbent group keeps a rather constant size of around 24 employees throughout

the period, the entrant group grows rapidly from 9 to 14 employees. Figure 1 depicts the ratio of the average number of employees of the entrant group to that of the incumbent group

$$E\left[L_t^{inc}\right] / E\left[L_t^{ent}\right] = \exp\left(\alpha_{ent,t}\right), \qquad (5)$$

together with the confidence interval of two standard errors.[2] The figure shows that the entrant firms on average start at 38 percent of the incumbent firms in terms of size, and then grow to 60 percent of the incumbent group after eight years.

[Table 6: Development of firm size]

[Figure 1: Ratio of the average firm size of the entrants to that of the incumbents]

This result is consistent with the previous findings that new firms are much smaller in size compared to old firms (e.g., Caves, 1998). Besides, it shows that the gap remains eight years after entry although the entrant firms grow rapidly directly after entry.

So far, only the average growth in size has been analyzed. The growth rate may, however, differ across different size groups. Thus, we will now compare size growth conditional on initial firm size. We define the initial size of each firm, L^0, as the number of employees in 1998 for both the entrant and incumbent groups. Adding cross terms with $\ln L^0$ to Eq.(2), we use the following model:

$$\ln L_{i,t} = \sum_t \alpha_{0,t} D_{i,t}^{year} + \sum_t \alpha_{ent,t} D_{i,t}^{ent} D_{i,t}^{year}$$
$$+ \sum_t \beta_{0,t} D_{i,t}^{year} \ln L_i^0 + \sum_t \beta_{ent,t} D_{i,t}^{ent} D_{i,t}^{year} \ln L_i^0 + \lambda_i^s + \epsilon_{i,t}. \quad (6)$$

The average number of employees of the incumbents with initial size L^0 is expressed by

$$E\left[L_t^{inc}|L^0\right] = \exp\left(\alpha_{0,t} + \beta_{0,t} \ln L^0 + \sigma_\epsilon^2/2\right), \qquad (7)$$

while that of the entrants with the same initial size is expressed by

$$E\left[L_t^{ent}|L^0\right] = \exp\left[\left(\alpha_{0,t} + \alpha_{ent,t}\right) + \left(\beta_{0,t} + \beta_{ent,t}\right) \ln L^0 + \sigma_\epsilon^2/2\right]. \qquad (8)$$

---

[2] The confidence interval is expressed by $\left[\exp\left(\alpha_{ent,t} - 2\sigma_{\alpha_{ent,t}}\right), \exp\left(\alpha_{ent,t} + 2\sigma_{\alpha_{ent,t}}\right)\right]$, where $\sigma_{\alpha_{ent,t}}$ denotes the standard error of $\alpha_{ent,t}$.

Table 7 presents $E\left[L_t^{inc}|L^0\right]$ and $E\left[L_t^{ent}|L^0\right]$ for the given initial sizes $L^0 = 1, 5, 10, 20$, and 50. There is a clear trend that, whereas the size of incumbents stays almost constant or slightly decreases, the entrants grow rapidly in the years after entry. The growth rate is high when the initial size is small. Figures 2–6 depict the ratios of the average number of employees of the entrant group to that of the incumbent group for different initial sizes,

$$E\left[L_t^{ent}|L^0\right]/E\left[L_t^{inc}|L^0\right] = \exp\left(\alpha_{ent,t} + \beta_{ent,t}\ln L^0\right), \qquad (9)$$

together with the confidence interval of two standard errors.

[Table 7: Development of firm size conditional on the initial size]

[Figures 2–6: Ratios of the average firm size of entrants to that of incumbents conditional on initial size]

New firms that started with only one employee are on average 93 percent larger in 2006 compared to incumbent firms with the same initial size, and the difference is statistically significant. The size difference between the two groups, however, decreases with initial size. With an initial size of 20 employees, the entrants are on average 11 percent larger in 2006. With an initial size of 50 employees, there is no significant difference.

These findings suggest that the size growth rate is negatively related to initial size, which is consistent with previous studies (e.g., Audretsch and Mahmood, 1994; Hart and Oulton, 1996; Audretsch et al., 1999a, 1999b; Heshmati, 2001).

5.2   Total factor productivity (TFP)

A similar analysis is implemented for TFP. Table 8 shows the average development of TFP for the two groups using Eq.(2), where the dependent variable $\ln L_{i,t}$ is replaced by the logarithm of TFP, $\ln TFP_{i,t}$. Figure 7 illustrates the ratio of the average TFP of the entrant group to that of the incumbent group, i.e., $E\left[TFP_t^{inc}\right]/E\left[TFP_t^{ent}\right]$ with the confidence interval of two standard errors. As shown, the entrant group has a significant drawback in the first three years. However, the gap disappears already in the fourth year. Note that the productivity of the entrant group in the first year after entry can be difficult to interpret, as some of the firms exist and operate only for a few months.

[Table 8: Development of labor productivity]

[Figure 7: Ratio of average TFP of the entrants to that of the incumbents]

We then analyze the productivity development conditional on initial size. Similar to firm size growth, the growth rate of labor productivity may differ across different size groups. Table 9 reports $E\left[TFP_t^{inc}|L^0\right]$ and $E\left[TFP_t^{ent}|L^0\right]$ for the given initial size $L^0 = 1$, 5, 10, 20, and 50. They are obtained using the equation Eq.(6) with $\ln L_{i,t}$ replaced by $\ln TFP_{i,t}$. Figures 8–12 illustrate the ratios of average TFP of entrants to that of incumbents conditional on initial size, i.e., $E\left[TFP_t^{inc}|L^0\right]/E\left[TFP_t^{ent}|L^0\right]$ with the confidence interval of two standard errors. As shown, new firms starting with one employee have no significant gap in labor productivity already one year after entry and gain a lead over the incumbent firms with the same initial size after four years. The gap reaches its maximum in the eighth year, where the entrants on average operate 14 percent more productively than the incumbents. For entrants with a larger initial size, it takes a longer time to catch up with the incumbents with the same initial size. With the initial sizes of $L^0 = 10$ and 20, it takes four years. With the initial size of $L^0 = 50$, the entrants never catch up with the incumbents with the same initial size.

[Table 9: Development of labor productivity, conditional on initial size]

[Figures 8–12: Ratios of average TFP of the entrants to that of the incumbents conditional on the initial size]

6   Conclusion

In this study, we investigate the relationship between the initial firm size and post-entry growth in size and total factor productivity (TFP). We use Swedish firm-level register-based data for the mining and manufacturing industries during the period 1997–2006 and compare two groups of firms: those that entered the market in 1998, which we call the "entrant group," and those that had been in business for at least 10 years in 1998.

The results show that the entrant firms were on average 38 percent of the average size of the incumbent firms in 1998, yet the relative average size of the entrants grew rapidly and reached 60 percent in 2006. However, the growth in size differs considerably depending on firm's initial size. While the new firms with 1–5 initial employees grew on average by 40–80 percent more than the

incumbents with similar initial size in eight years, the new firms with 20 initial employees grew by only 13 percent more than the incumbents. The relationship between initial firm size and subsequent change in relative position is therefore negative.

Regarding TFP, the average TFP of all firms in the entrant group was about 15 percent lower than the average of the incumbent group in 1998, yet the difference disappeared after three years. However, we find a large difference in growth pattern depending on initial firm size. The entrants with one initial employee caught up with the incumbents of similar initial size already in the second year, and gained a lead in the fifth year. It takes three years for the productivity gap between the entrants with 10 initial employees and the incumbents with the same size to disappear. The entrants with more than 20 initial employees never caught up with the entrants of the same size during the nine years analyzed in this study.

The current study shows that a negative relationship is not only found between initial firm size and post-entry growth in size, but also between initial firm size and post-entry productivity development.

References

[1] Audretsch, David B. (1995), *Innovation and Industry Evolution*, Cambridge, Mass.: MIT Press.

[2] Audretsch, David B.; Klomp, Luuk; and Thurik, A. Roy (1999a). "Innovation, Industry Evolution, and Employment", In D. Audretsch and R. Thurik (eds.), *Nordic Labour Market Research on Register Data*, 230-52, Cambridge University Press.

[3] Audretsch, David B.; Santarelli, Enrico; and Vivarelli, Marco (1999b), "Start-up size and industrial dynamics: some evidence from Italian manufacturing," *International Journal of Industrial Organization*, 17(7), 965-83.

[4] Audretsch, David B. and Mahmood, Talat (1994), "Firm Selection and Industry Evolution: The Post-entry Performance of New Firms," *Journal of Evolutionary Economics*, 4(3), 243-60.

[5] Aw, Bee Yan; Chen, Xiaomin; and Roberts, Mark J. (2001), "Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing," *Journal of Development Economics*, 66(1), 51-86.

[6] Baldwin, John R. and Rafiquzzaman, Mohammed (1995), "Selection versus evolutionary adaptation: Learning and post-entry performance," *International Journal of Industrial Organization*, 13(4), 501-22.

[7] Bellone, Flora; Musso, Patrick; Quéré, Michel; and Nesta, Lionel (2006), "Productivity and Market Selection of French Manufacturing Firms in the Nineties," *Revue de l'OFCE, Presses de Sciences-Po*, 97(5), 319-49.

[8] Bellone, Flora; Musso, Patrick; Quéré, Michel; and Nesta, Lionel (2008), "Market selection along the firm life cycle," *Industrial and Corporate Change*, 17(4), 753-77.

[9] Caves, Douglas W.; Christensen, Laurits R.; and Diewert, W Erwin (1982), "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers," *Economic Journal*, 92(365), 73-86.

[10] de Kok, Jan; Brouwer, Peter; and Fris, Pieter (2005), "Can firm age account for productivity differences?," Scales Research Reports N200421, EIM Business and Policy Research.

[11] Dunne, Timothy; Roberts, Mark J.; and Samuelson, Larry (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, 19(4), 495-515.

[12] Good, David H.; Nadiri, M. Ishaq; and Sickles, Robin C. (1996), "Index Number and Factor Demand Approaches to the Estimation of Productivity," NBER Working Papers 5790, National Bureau of Economic Research.

[13] Hanson, Andrew (2009), "Local employment, poverty, and property value effects of geographically-targeted tax incentives: An instrumental variables approach," *Regional Science and Urban Economics*, 39(6), 721-31.

[14] Hart, Peter E. and Oulton, Nicholas (1996), "Growth and size of firms," *The Economic Journal*, 106(438), 1242–52.

[15] Heshmati, Almas (2001), "On the Growth of Micro and Small Firms: Evidence from Sweden," *Small Business Economics*, 17(3), 213-28.

[16] Jensen, J. Bradford; McGuckin, Robert H.; and Stiroh, Kevin J. (2001), "The impact of vintage and survival on productivity: Evidence from cohorts of U.S. manufacturing plants," *The Review of Economics and Statistics*, 83(2), 323-32.

[17] Kolko, Jed and Neumark, David (2009), "Do Some Enterprise Zones Create Jobs?," NBER Working Papers 15206, National Bureau of Economic Research.

[18] Neumark, David and Kolko, Jed (2010), "Do enterprise zones create jobs? Evidence from California's enterprise zone program," *Journal of Urban Economics*, 68(1), 1-19

[19] Persson, Helena (2004), "The Survival and Growth of New Establishments in Sweden, 1987-1995," *Small Business Economics*, 23(5), 423-40.

[20] Power, Laura (1998), "The missing link: technology, investment and productivity," *Review of Economics and Statistics*, 80(2), 300-13.

[21] SOU (Statens offentliga utredningar) (2012), "Nystartszoner", SOU 2012:50.

						labor share of	
year	all firms	entry firms	exit firms	entry rate	exit rate	entry firms	exit firms
1997	17566		1407		0.080		0.035
1998	17329	1170	1358	0.068	0.078	0.028	0.055
1999	17218	1247	1091	0.072	0.063	0.050	0.083
2000	17491	1364	1081	0.078	0.062	0.044	0.040
2001	17568	1158	1084	0.066	0.062	0.037	0.071
2002	17510	1026	989	0.059	0.056	0.025	0.040
2003	17717	1196	845	0.068	0.048	0.036	0.039
2004	18061	1189	797	0.066	0.044	0.028	0.033
2005	18334	1070	806	0.058	0.044	0.020	0.030
2006	18855	1327		0.070		0.023	

[Table 1: Entry and exit in Swedish mining and manufacturing industries]

year	all firms	entry firms	exit firms
1997	33.9		15.0
1998	35.2	14.5	24.9
1999	35.1	24.0	45.7
2000	34.7	19.5	22.5
2001	34.6	19.6	40.0
2002	32.8	14.1	23.1
2003	32.0	17.2	25.9
2004	31.2	13.1	23.4
2005	30.4	10.5	20.7
2006	29.4	9.5	

[Table 2: Average size (in terms of employees) of all firms, entry firms and exit firms by year]

NACE 2-digit industry classification	entry rate	exit rate	average firm size (the number of employees)		
			all firms	entry firms	exit firms
10-14 mining and quarrying	0.051	0.050	36.0	27.9	37.8
15-16 food products	0.082	0.067	39.7	21.7	35.8
17-19 textiles	0.052	0.063	16.7	6.3	20.1
20 wood products	0.063	0.055	18.6	11.1	16.4
21 pulp and paper	0.054	0.047	158.2	68.4	70.8
22 publishing and printing	0.061	0.070	15.5	8.4	10.9
23 coke, petroleum and other fuels	0.076	0.076	122.8	53.2	88.4
24 chemicals	0.069	0.057	103.5	37.9	34.7
25 rubber and plastic products	0.047	0.044	25.6	14.8	20.4
26 other non-metallic mineral products	0.045	0.048	37.0	31.1	30.3
27 basic metals	0.055	0.036	146.7	20.9	39.4
28 fabricated metal products	0.057	0.043	15.9	7.4	16.5
29 machinery	0.056	0.046	41.3	17.0	22.9
30 office machinery and computers	0.060	0.079	26.6	11.1	11.3
31 electrical machinery	0.051	0.053	38.6	34.8	83.6
32 communication equipment	0.095	0.076	53.7	39.5	102.0
33 medical and optical instruments	0.053	0.050	23.5	13.9	21.3
34 motor vehicles	0.062	0.048	171.5	56.5	115.1
35 other transport equipment	0.060	0.047	37.8	29.1	47.7
36-37 furniture, recycling and else	0.078	0.062	23.5	8.3	25.1
TOTAL	0.060	0.053	32.9	15.9	26.7

[Table 3: Entry rates, exit rates and average firm size by industry] (? kommer att fyllas)

(a) The number of survived firms

entry year	1998	1999	2000	2001	2002	2003	2004	2005	2006
				year					
1998	1170	976	870	778	715	666	645	618	589
1999		1247	1097	976	887	817	772	725	687
2000			1364	1206	1072	976	923	876	837
2001				1158	1038	946	890	837	798
2002					1026	921	851	797	754
2003						1196	1065	986	922
2004							1189	1093	1008
2005								1070	972
2006									1327

(b) The survival rate

entry year	1998	1999	2000	2001	2002	2003	2004	2005	2006
				year					
1998	100.0	83.4	74.4	66.5	61.1	56.9	55.1	52.8	50.3
1999		100.0	88.0	78.3	71.1	65.5	61.9	58.1	55.1
2000			100.0	88.4	78.6	71.6	67.7	64.2	61.4
2001				100.0	89.6	81.7	76.9	72.3	68.9
2002					100.0	89.8	82.9	77.7	73.5
2003						100.0	89.0	82.4	77.1
2004							100.0	91.9	84.8
2005								100.0	90.8
2006									100.0

[Table 4: Cohort survival]

the incumbent group					
year	Obs.	Mean	Std. Dev.	Min	Max
(a) the number of employees					
1998	8,862	45	262	1	9,681
2006	6,306	46	324	1	18839
(b) log of labor productivity					
1998	8,862	-1.004	0.508	-5.383	3.331
2006	6,306	-0.917	0.592	-6.354	3.743
(c) TFP					
1998	8,862	-0.828	0.452	-5.615	2.316
2006	6,306	-0.858	0.539	-5.264	2.549

the entrant group					
year	Obs.	Mean	Std. Dev.	Min	Max
(a) the number of employees					
1998	1170	14	72	1	1809
2006	589	16	64	1	1120
(b) log of labor productivity					
1998	1170	-1.186	0.701	-5.495	2.788
2006	589	-0.941	0.644	-5.761	2.392
(c) TFP					
1998	1170	-0.985	0.642	-5.195	2.758
2006	589	-0.868	0.591	-4.878	2.502

[Table 5: Comparison of two groups]

	year								
	1998	1999	2000	2001	2002	2003	2004	2005	2006
Incumbent group	23.5	23.6	23.9	24.1	23.9	23.7	23.2	23.2	23.1
Entrant group	8.9	11.5	12.5	12.9	13.0	13.2	13.2	13.6	13.8

[Table 6 Development of firm size (the number of employees)]

	year								
	1998	1999	2000	2001	2002	2003	2004	2005	2006
Initial size: L = 1									
Incumbent group	1.13	1.17	1.17	1.19	1.19	1.19	1.17	1.18	1.17
Entrant group	1.12	1.63	1.83	1.88	1.95	2.06	2.06	2.15	2.26
Initial size: L = 5									
Incumbent group	5.60	5.56	5.59	5.65	5.62	5.56	5.47	5.48	5.46
Entrant group	5.58	6.91	7.39	7.58	7.67	7.80	7.72	7.92	7.94
Initial size: L = 10									
Incumbent group	11.16	10.89	10.98	11.03	10.94	10.81	10.63	10.63	10.83
Entrant group	11.12	12.89	13.47	13.80	13.83	13.85	13.65	13.87	13.77
Initial size: L = 20									
Incumbent group	22.23	21.32	21.55	21.53	21.33	21.01	20.67	20.61	20.54
Entrant group	22.17	24.04	24.56	25.13	24.95	24.59	24.13	24.30	23.42
Initial size: L = 50									
Incumbent group	55.27	51.83	52.56	52.16	51.53	50.59	49.77	49.48	49.34
Entrant group	55.20	54.79	54.36	55.52	54.38	52.51	51.25	51.00	47.90

[Table 7 Development of firm size (the number of employees) conditional on the initial size]

year	1998	1999	2000	2001	2002	2003	2004	2005	2006
Incumbent group	0.928	0.942	0.983	0.923	0.905	0.902	0.909	0.891	0.902
Entrant group	0.793	0.885	0.911	0.910	0.920	0.881	0.905	0.905	0.897

[Table 8 Development of TFP]

year	1998	1999	2000	2001	2002	2003	2004	2005	2006
Initial size: L = 1									
Incumbent group	0.438	0.436	0.445	0.425	0.404	0.394	0.395	0.377	0.379
Entrant group	0.387	0.436	0.448	0.455	0.446	0.420	0.441	0.431	0.413
Initial size: L = 5									
Incumbent group	0.467	0.471	0.488	0.461	0.447	0.443	0.446	0.434	0.438
Entrant group	0.411	0.452	0.457	0.460	0.457	0.433	0.442	0.436	0.432
Initial size: L = 10									
Incumbent group	0.480	0.487	0.508	0.477	0.468	0.466	0.470	0.461	10.969
Entrant group	0.422	0.459	0.461	0.462	0.462	0.439	0.443	0.438	13.947
Initial size: L = 20									
Incumbent group	0.493	0.503	0.529	0.494	0.489	0.491	0.495	0.489	0.496
Entrant group	0.433	0.466	0.465	0.465	0.467	0.445	0.443	0.439	0.448
Initial size: L = 50									
Incumbent group	0.511	0.525	0.557	0.518	0.519	0.525	0.531	0.529	0.539
Entrant group	0.447	0.476	0.470	0.468	0.474	0.453	0.444	0.442	0.459

[Table 9 Development of TFP conditional on the initial size]
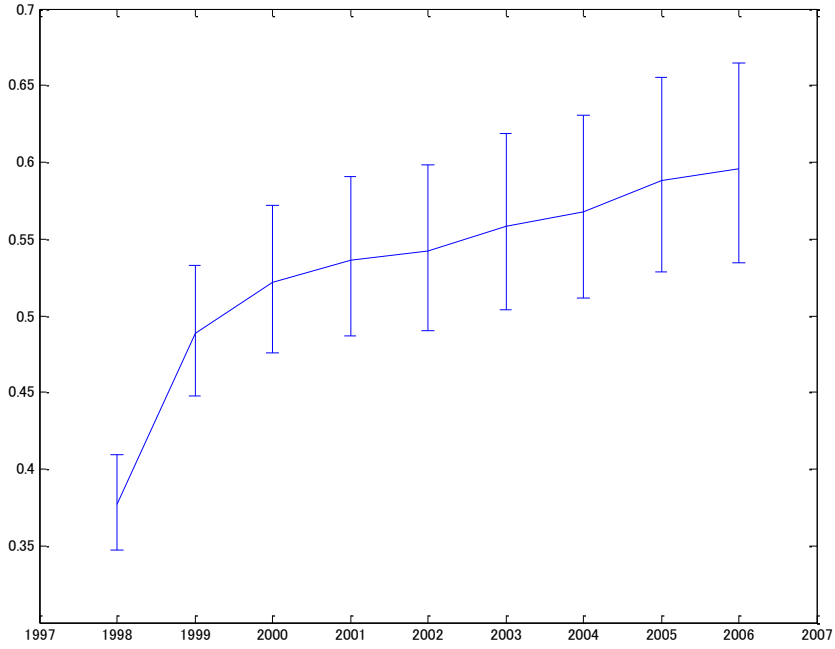
[Figure 1: Ratio of the average firm size of the entrants to that of the incumbents]
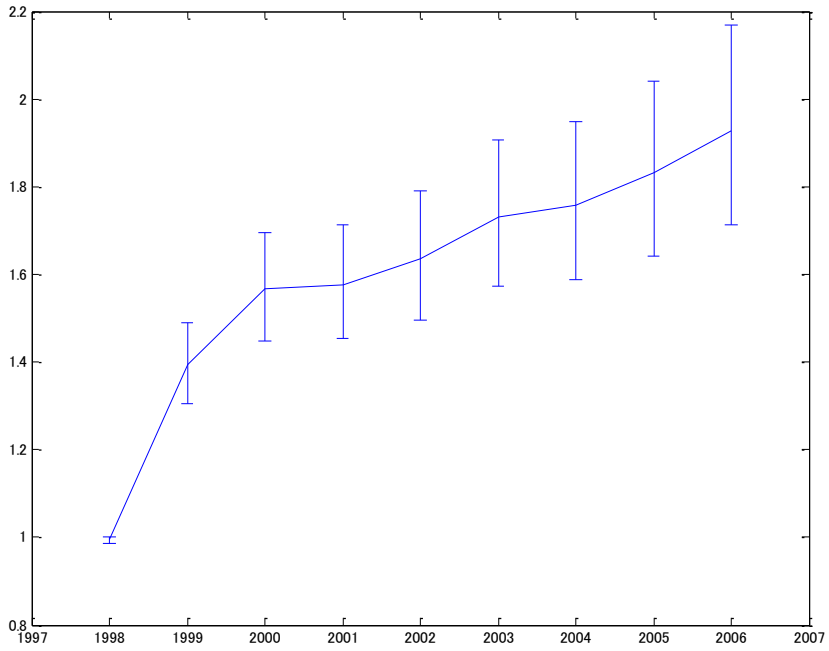
[Figure 2: Ratio of the average firm size of the entrants to that of the incumbents conditional on the initial size $L^0 = 1$]
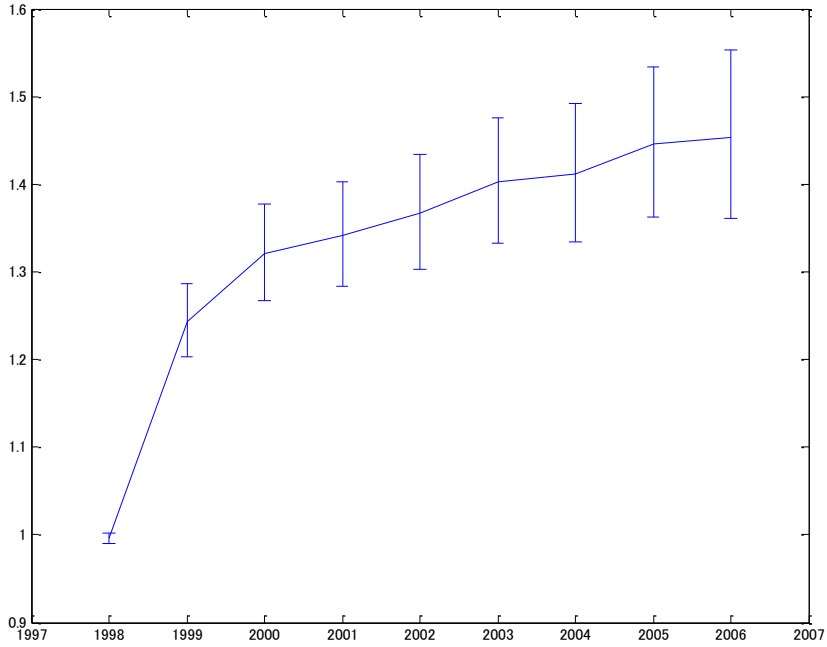
[Figure 3: Ratio of the average firm size of the entrants to that of the incumbents conditional on the initial size $L^0 = 5$]
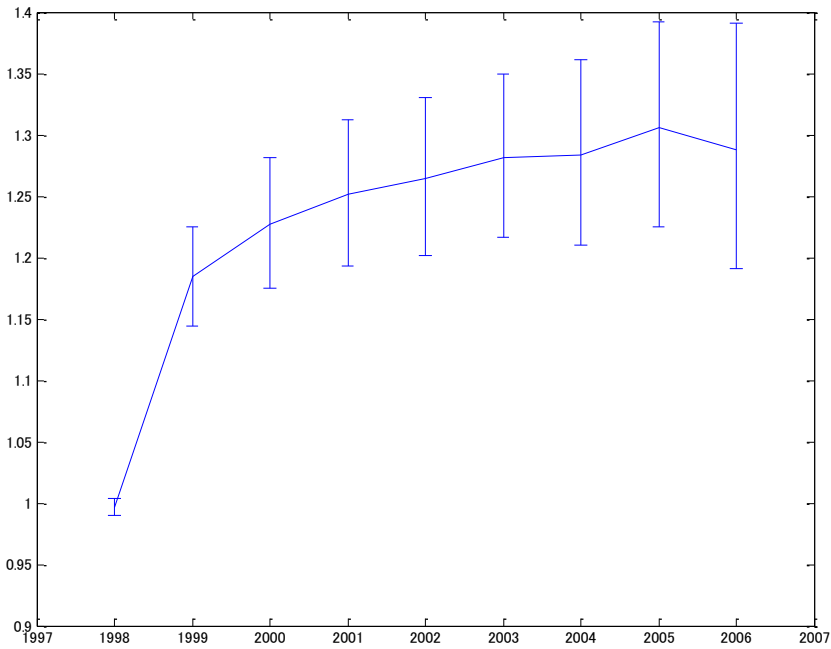
[Figure 4: Ratio of the average firm size of the entrants to that of the incumbents conditional on the initial size $L^0 = 10$]
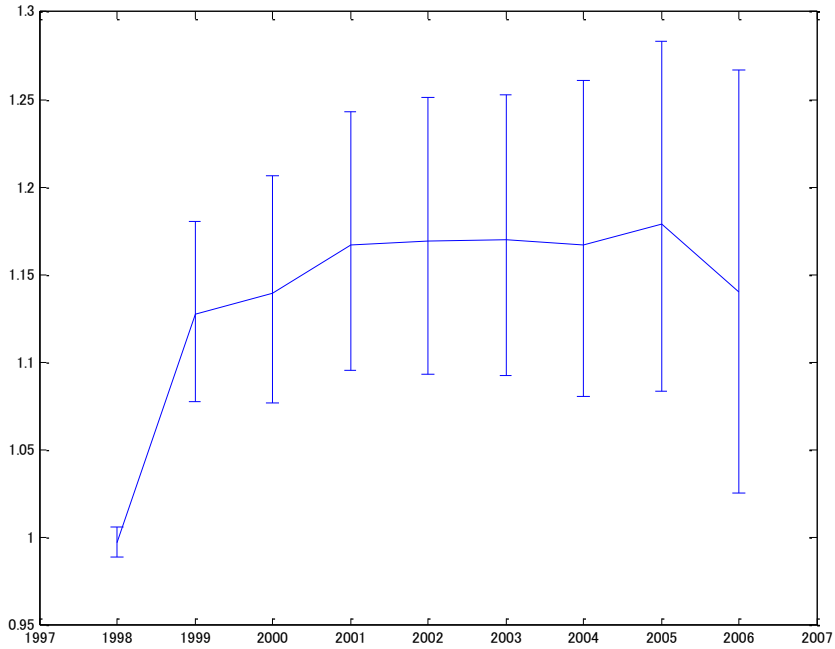
[Figure 5: Ratio of the average firm size of the entrants to that of the incumbents conditional on the initial size $L^0 = 20$]´
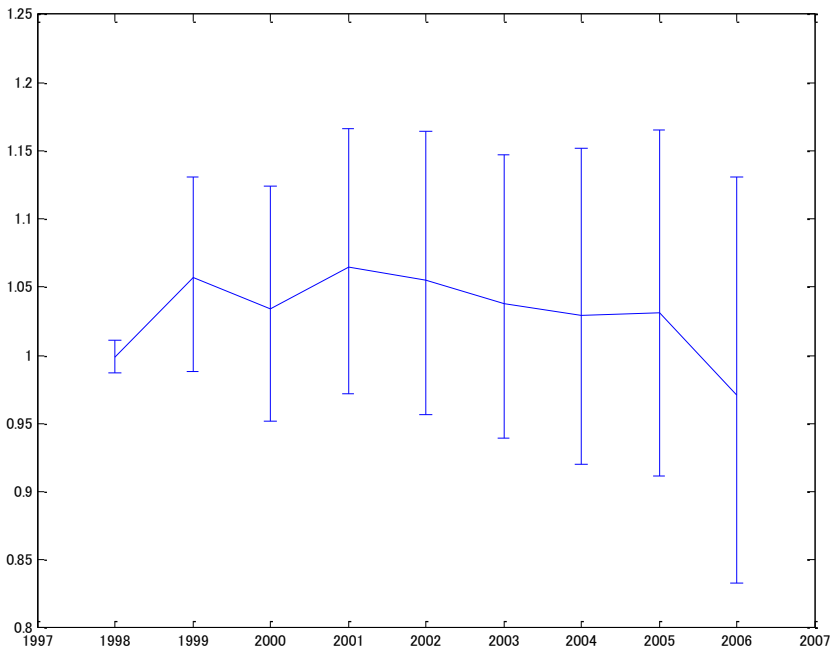
[Figure 6: Ratio of the average firm size of the entrants to that of the incumbents conditional on the initial size $L^0 = 50$]´
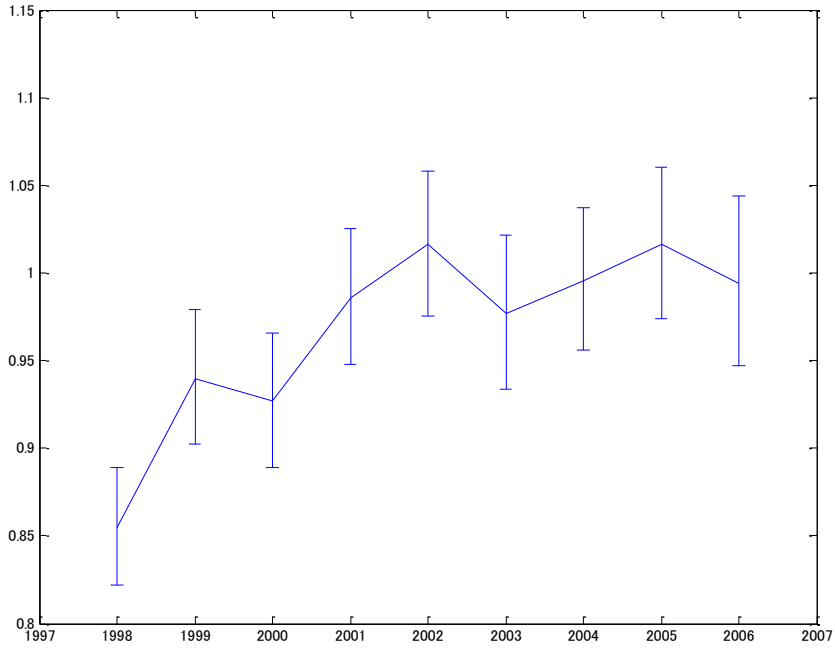
[Figure 7: Ratio of the average TFP of the entrants to that of the incumbents]
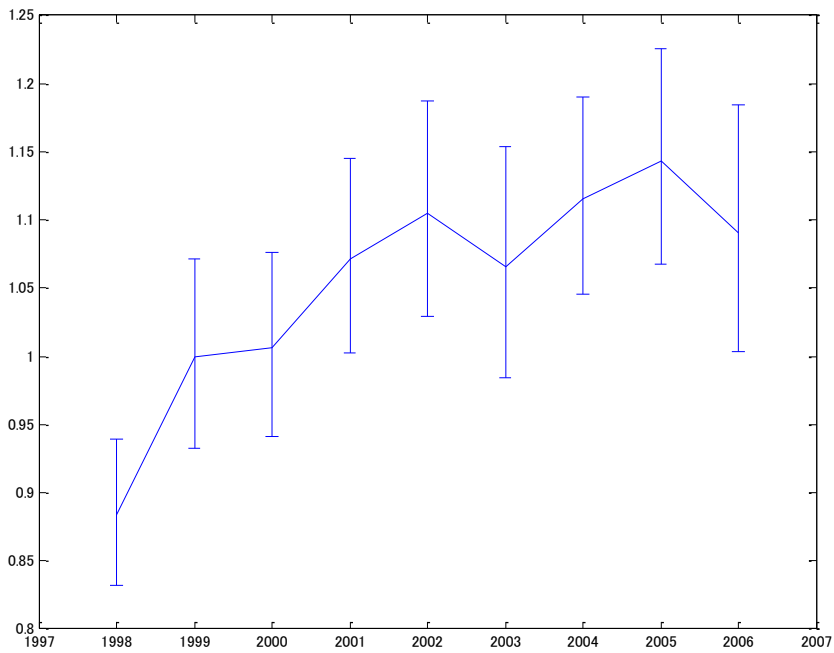
[Figure 8: Ratio of average TFP of the entrants to that of the incumbents conditional on the initial size $L^0 = 1$]
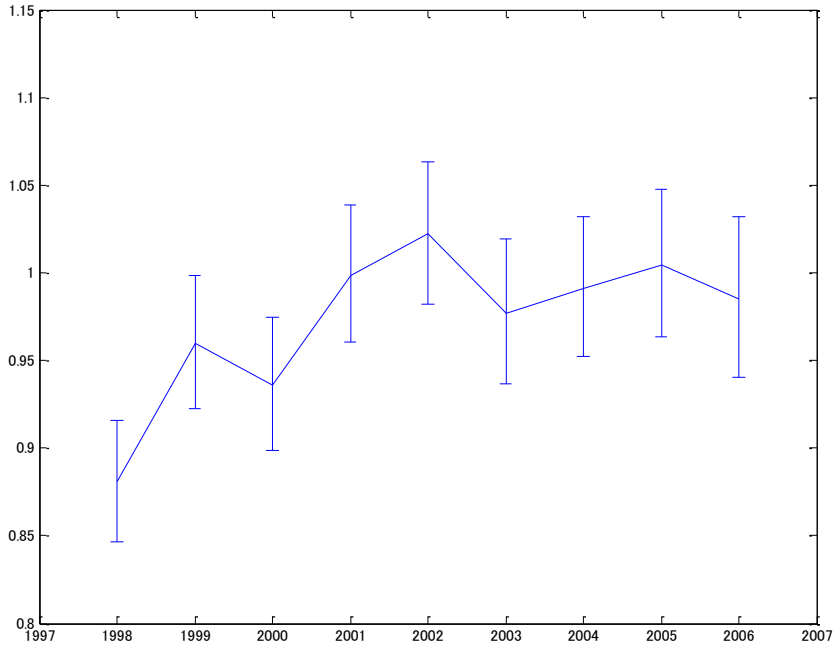
[Figure 9: Ratio of average TFP of the entrants to that of the incumbents conditional on the initial size $L^0 = 5$]
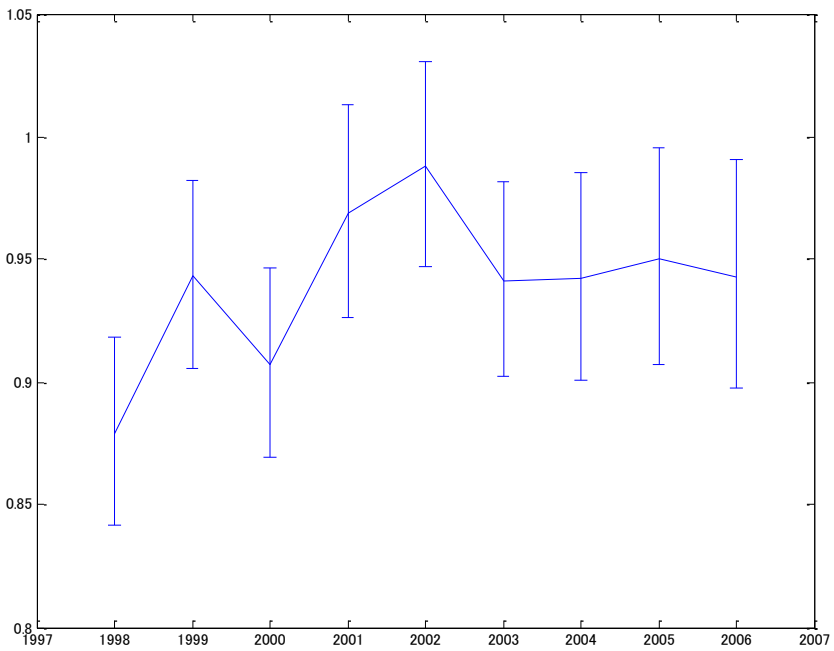
[Figure 10: Ratio of average TFP of the entrants to that of the incumbents conditional on the initial size $L^0 = 10$]
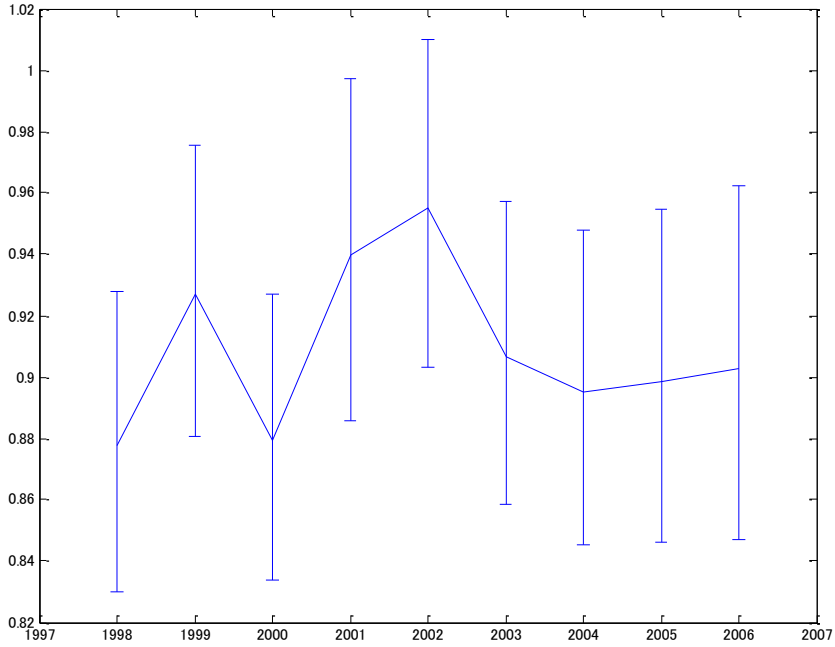
[Figure 11: Ratio of average TFP of the entrants to that of the incumbents conditional on the initial size $L^0 = 20$]
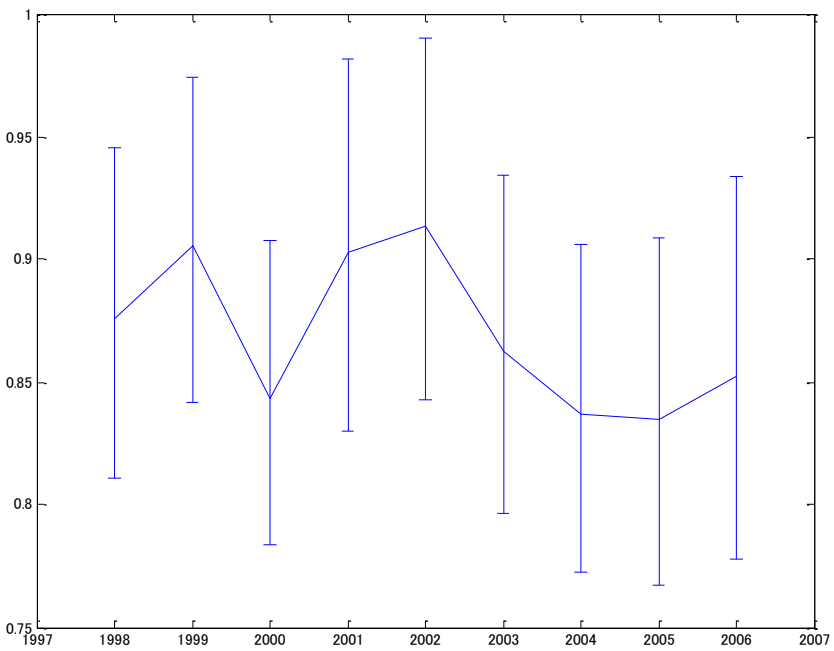
[Figure 12: Ratio of average TFP of the entrants to that of the incumbents conditional on the initial size $L^0 = 50$]

Previous doctoral theses in the Department of Economics, Gothenburg

Avhandlingar publicerade innan serien Ekonomiska Studier startades
(Theses published before the series Ekonomiska Studier was started):

Östman, Hugo (1911), Norrlands ekonomiska utveckling
Moritz, Marcus (1911), Den svenska tobaksindustrien
Sundbom, I. (1933), Prisbildning och ändamålsenlighet
Gerhard, I. (1948), Problem rörande Sveriges utrikeshandel 1936/38
Hegeland, Hugo (1951), The Quantity Theory of Money
Mattsson, Bengt (1970), Cost-Benefit analys
Rosengren, Björn (1975), Valutareglering och nationell ekonomisk politik
Hjalmarsson, Lennart (1975), Studies in a Dynamic Theory of Production and its Applications
Örtendahl, Per-Anders (1975), Substitutionsaspekter på produktionsprocessen vid massaframställning
Anderson, Arne M. (1976), Produktion, kapacitet och kostnader vid ett helautomatiskt emballageglasbruk
Ohlsson, Olle (1976), Substitution och odelbarheter i produktionsprocessen vid massaframställning
Gunnarsson, Jan (1976), Produktionssystem och tätortshierarki – om sambandet mellan rumslig och ekonomisk struktur
Köstner, Evert (1976), Optimal allokering av tid mellan utbildning och arbete
Wigren, Rune (1976), Analys av regionala effektivitetsskillnader inom industribranscher
Wästlund, Jan (1976), Skattning och analys av regionala effektivitetsskillnader inom industribranscher
Flöjstad, Gunnar (1976), Studies in Distortions, Trade and Allocation Problems
Sandelin, Bo (1977), Prisutveckling och kapitalvinster på bostadsfastigheter
Dahlberg, Lars (1977), Empirical Studies in Public Planning
Lönnroth, Johan (1977), Marxism som matematisk ekonomi
Johansson, Börje (1978), Contributions to Sequential Analysis of Oligopolistic Competition

Ekonomiska Studier, utgivna av Nationalekonomiska institutionen vid Göteborgs Universitet. Nr 1 och 4 var inte doktorsavhandlingar. (The contributions to the department series 'Ekonomiska Studier' where no. 1 and 4 were no doctoral theses):

2. **Ambjörn, Erik** (1959), Svenskt importberoende 1926-1956: en ekonomisk-statistisk kartläggning med kommentarer
3. **Landgren, K-G.** (1960), Den "Nya ekonomien" i Sverige: J.M. Keynes, E. Wigfors och utecklingen 1927-39
5. **Bigsten, Arne** (1979), Regional Inequality and Development: A Case Study of Kenya
6. **Andersson, Lars** (1979), Statens styrning av de kommunala budgetarnas struktur (Central Government Influence on the Structure of the Municipal Budget)
7. **Gustafsson, Björn** (1979), Inkomst- och uppväxtförhållanden (Income and Family Background)

8. **Granholm, Arne** (1981), Interregional Planning Models for the Allocation of Private and Public Investments
9. **Lundborg, Per** (1982), Trade Policy and Development: Income Distributional Effects in the Less Developed Countries of the US and EEC Policies for Agricultural Commodities
10. **Juås, Birgitta** (1982), Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. (Valuation of Personal Injuries. A comparison of Explicit and Implicit Values)
11. **Bergendahl, Per-Anders** (1982), Energi och ekonomi - tillämpningar av input-output analys (Energy and the Economy - Applications of Input-Output Analysis)
12. **Blomström, Magnus** (1983), Foreign Investment, Technical Efficiency and Structural Change - Evidence from the Mexican Manufacturing Industry
13. **Larsson, Lars-Göran** (1983), Comparative Statics on the Basis of Optimization Methods
14. **Persson, Håkan** (1983), Theory and Applications of Multisectoral Growth Models
15. **Sterner, Thomas** (1986), Energy Use in Mexican Industry.
16. **Flood, Lennart** (1986), On the Application of Time Use and Expenditure Allocation Models.
17. **Schuller, Bernd-Joachim** (1986), Ekonomi och kriminalitet - en empirisk undersökning av brottsligheten i Sverige (Economics of crime - an empirical analysis of crime in Sweden)
18. **Walfridson, Bo** (1987), Dynamic Models of Factor Demand. An Application to Swedish Industry.
19. **Stålhammar, Nils-Olov** (1987), Strukturomvandling, företagsbeteende och förväntningsbildning inom den svenska tillverkningsindustrin (Structural Change, Firm Behaviour and Expectation Formation in Swedish Manufactury)
20. **Anxo, Dominique** (1988), Sysselsättningseffekter av en allmän arbetstidsför-kortning (Employment effects of a general shortage of the working time)
21. **Mbelle, Ammon** (1988), Foreign Exchange and Industrial Development: A Study of Tanzania.
22. **Ongaro, Wilfred** (1988), Adoption of New Farming Technology: A Case Study of Maize Production in Western Kenya.
23. **Zejan, Mario** (1988), Studies in the Behavior of Swedish Multinationals.
24. **Görling, Anders** (1988), Ekonomisk tillväxt och miljö. Förorenings-struktur och ekonomiska effekter av olika miljövårdsprogram. (Economic Growth and Environment. Pollution Structure and Economic Effects of Some Environmental Programs).
25. **Aguilar, Renato** (1988), Efficiency in Production: Theory and an Application on Kenyan Smallholders.
26. **Kayizzi-Mugerwa, Steve** (1988), External Shocks and Adjustment in Zambia.

27. **Bornmalm-Jardelöw, Gunilla** (1988), Högre utbildning och arbetsmarknad (Higher Education and the Labour Market)
28. **Tansini, Ruben** (1989), Technology Transfer: Dairy Industries in Sweden and Uruguay.
29. **Andersson, Irene** (1989), Familjebeskattning, konsumtion och arbetsutbud - En ekonometrisk analys av löne- och inkomstelasticiteter samt policysimuleringar för svenska hushåll (Family Taxation, Consumption and Labour Supply - An Econometric Analysis of Wage and Income Elasticities and Policy Simulations for Swedish Households)
30. **Henrekson, Magnus** (1990), An Economic Analysis of Swedish Government Expenditure
31. **Sjöö, Boo** (1990), Monetary Policy in a Continuous Time Dynamic Model for Sweden
32. **Rosén, Åsa** (1991), Contributions to the Theory of Labour Contracts.
33. **Loureiro, Joao M. de Matos** (1992), Foreign Exchange Intervention, Sterilization and Credibility in the EMS: An Empirical Study
34. **Irandoust, Manuchehr** (1993), Essays on the Behavior and Performance of the Car Industry
35. **Tasiran, Ali Cevat** (1993), Wage and Income Effects on the Timing and Spacing of Births in Sweden and the United States
36. **Milopoulos, Christos** (1993), Investment Behaviour under Uncertainty: An Econometric Analysis of Swedish Panel Data
37. **Andersson, Per-Åke** (1993), Labour Market Structure in a Controlled Economy: The Case of Zambia
38. **Storrie, Donald W.** (1993), The Anatomy of a Large Swedish Plant Closure
39. **Semboja, Haji Hatibu Haji** (1993), Energy and Development in Kenya
40. **Makonnen, Negatu** (1993), Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho
41. **Julin, Eva** (1993), Structural Change in Rural Kenya
42. **Durevall, Dick** (1993), Essays on Chronic Inflation: The Brazilian Experience
43. **Veiderpass, Ann** (1993), Swedish Retail Electricity Distribution: A Non-Parametric Approach to Efficiency and Productivity Change
44. **Odeck, James** (1993), Measuring Productivity Growth and Efficiency with Data Envelopment Analysis: An Application on the Norwegian Road Sector
45. **Mwenda, Abraham** (1993), Credit Rationing and Investment Behaviour under Market Imperfections: Evidence from Commercial Agriculture in Zambia
46. **Mlambo, Kupukile** (1993), Total Factor Productivity Growth: An Empirical Analysis of Zimbabwe's Manufacturing Sector Based on Factor Demand Modelling
47. **Ndung'u, Njuguna** (1993), Dynamics of the Inflationary Process in Kenya
48. **Modén, Karl-Markus** (1993), Tax Incentives of Corporate Mergers and Foreign Direct Investments
49. **Franzén, Mikael** (1994), Gasoline Demand - A Comparison of Models
50. **Heshmati, Almas** (1994), Estimating Technical Efficiency, Productivity Growth And Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture

51. **Salas, Osvaldo** (1994), Efficiency and Productivity Change: A Micro Data Case Study of the Colombian Cement Industry
52. **Bjurek, Hans** (1994), Essays on Efficiency and Productivity Change with Applications to Public Service Production
53. **Cabezas Vega, Luis** (1994), Factor Substitution, Capacity Utilization and Total Factor Productivity Growth in the Peruvian Manufacturing Industry
54. **Katz, Katarina** (1994), Gender Differentiation and Discrimination. A Study of Soviet Wages
55. **Asal, Maher** (1995), Real Exchange Rate Determination and the Adjustment Process: An Empirical Study in the Cases of Sweden and Egypt
56. **Kjulin, Urban** (1995), Economic Perspectives on Child Care
57. **Andersson, Göran** (1995), Volatility Forecasting and Efficiency of the Swedish Call Options Market
58. **Forteza, Alvaro** (1996), Credibility, Inflation and Incentive Distortions in the Welfare State
59. **Locking, Håkan** (1996), Essays on Swedish Wage Formation
60. **Välilä, Timo** (1996), Essays on the Credibility of Central Bank Independence
61. **Yilma, Mulugeta** (1996), Measuring Smallholder Efficiency: Ugandan Coffee and Food-Crop Production
62. **Mabugu, Ramos E.** (1996), Tax Policy Analysis in Zimbabwe Applying General Equilibrium Models
63. **Johansson, Olof** (1996), Welfare, Externalities, and Taxation; Theory and Some Road Transport Applications.
64. **Chitiga, Margaret** (1996), Computable General Equilibrium Analysis of Income Distribution Policies in Zimbabwe
65. **Leander, Per** (1996), Foreign Exchange Market Behavior Expectations and Chaos
66. **Hansen, Jörgen** (1997), Essays on Earnings and Labor Supply
67. **Cotfas, Mihai** (1997), Essays on Productivity and Efficiency in the Romanian Cement Industry
68. **Horgby, Per-Johan** (1997), Essays on Sharing, Management and Evaluation of Health Risks
69. **Nafar, Nosratollah** (1997), Efficiency and Productivity in Iranian Manufacturing Industries
70. **Zheng, Jinghai** (1997), Essays on Industrial Structure, Technical Change, Employment Adjustment, and Technical Efficiency
71. **Isaksson, Anders** (1997), Essays on Financial Liberalisation in Developing Countries: Capital mobility, price stability, and savings
72. **Gerdin, Anders** (1997), On Productivity and Growth in Kenya, 1964-94
73. **Sharifi, Alimorad** (1998), The Electricity Supply Industry in Iran: Organization, performance and future development
74. **Zamanian, Max** (1997), Methods for Mutual Fund Portfolio Evaluation: An application to the Swedish market
75. **Manda, Damiano Kulundu** (1997), Labour Supply, Returns to Education, and the Effect of Firm Size on Wages: The case of Kenya
76. **Holmén, Martin** (1998), Essays on Corporate Acquisitions and Stock Market Introductions

77. **Pan, Kelvin** (1998), Essays on Enforcement in Money and Banking
78. **Rogat, Jorge** (1998), The Value of Improved Air Quality in Santiago de Chile
79. **Peterson, Stefan** (1998), Essays on Large Shareholders and Corporate Control
80. **Belhaj, Mohammed** (1998), Energy, Transportation and Urban Environment in Africa: The Case of Rabat-Salé, Morocco
81. **Mekonnen, Alemu** (1998), Rural Energy and Afforestation: Case Studies from Ethiopia
82. **Johansson, Anders** (1998), Empirical Essays on Financial and Real Investment Behavior
83. **Köhlin, Gunnar** (1998), The Value of Social Forestry in Orissa, India
84. **Levin, Jörgen** (1998), Structural Adjustment and Poverty: The Case of Kenya
85. **Ncube, Mkhululi** (1998), Analysis of Employment Behaviour in Zimbabwe
86. **Mwansa, Ladslous** (1998), Determinants of Inflation in Zambia
87. **Agnarsson, Sveinn** (1998), Of Men and Machines: Essays in Applied Labour and Production Economics
88. **Kadenge, Phineas** (1998), Essays on Macroeconomic Adjustment in Zimbabwe: Inflation, Money Demand, and the Real Exchange Rate
89. **Nyman, Håkan** (1998), An Economic Analysis of Lone Motherhood in Sweden
90. **Carlsson, Fredrik** (1999), Essays on Externalities and Transport
91. **Johansson, Mats** (1999), Empirical Studies of Income Distribution
92. **Alemu, Tekie** (1999), Land Tenure and Soil Conservation: Evidence from Ethiopia
93. **Lundvall, Karl** (1999), Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94
94. **Zhang, Jianhua** (1999), Essays on Emerging Market Finance
95. **Mlima, Aziz Ponary** (1999), Four Essays on Efficiency and Productivity in Swedish Banking
96. **Davidsen, Björn-Ivar** (2000), Bidrag til den økonomisk-metodologiske tenkningen (Contributions to the Economic Methodological Thinking)
97. **Ericson, Peter** (2000), Essays on Labor Supply
98. **Söderbom, Måns** (2000), Investment in African Manufacturing: A Microeconomic Analysis
99. **Höglund, Lena** (2000), Essays on Environmental Regulation with Applications to Sweden
100. **Olsson, Ola** (2000), Perspectives on Knowledge and Growth
101. **Meuller, Lars** (2000), Essays on Money and Credit
102. **Österberg, Torun** (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions
103. **Kalinda Mkenda, Beatrice** (2001), Essays on Purchasing Power Parity, RealExchange Rate, and Optimum Currency Areas
104. **Nerhagen, Lena** (2001), Travel Demand and Value of Time - Towards an Understanding of Individuals Choice Behavior

161. **Bezabih, Mintewab** (2007), Essays on Land Lease Markets, Productivity, Biodiversity, and Environmental Variability
162. **Visser, Martine** (2007), Fairness, Reciprocity and Inequality: Experimental Evidence from South Africa
163. **Holm, Louise** (2007), A Non-Stationary Perspective on the European and Swedish Business Cycle
164. **Herbertsson, Alexander** (2007), Pricing Portfolio Credit Derivatives
165. **Johansson, Anders C.** (2007), Essays in Empirical Finance: Volatility, Interdependencies, and Risk in Emerging Markets
166. **Ibáñez Díaz, Marcela** (2007), Social Dilemmas: The Role of Incentives, Norms and Institutions
167. **Ekbom, Anders** (2007), Economic Analysis of Soil Capital, Land Use and Agricultural Production in Kenya
168. **Sjöberg, Pål** (2007), Essays on Performance and Growth in Swedish Banking
169. **Palma Aguirre, Grisha Alexis** (2008), Explaining Earnings and Income Inequality in Chile
170. **Akay, Alpaslan** (2008), Essays on Microeconometrics and Immigrant Assimilation
171. **Carlsson, Evert** (2008), After Work – Investing for Retirement
172. **Munshi, Farzana** (2008), Essays on Globalization and Occupational Wages
173. **Tsakas, Elias** (2008), Essays on Epistemology and Evolutionary Game Theory
174. **Erlandzon, Karl** (2008), Retirement Planning: Portfolio Choice for Long-Term Investors
175. **Lampi, Elina** (2008), Individual Preferences, Choices, and Risk Perceptions – Survey Based Evidence
176. **Mitrut, Andreea** (2008), Four Essays on Interhousehold Transfers and Institutions in Post-Communist Romania
177. **Hansson, Gustav** (2008), Essays on Social Distance, Institutions, and Economic Growth
178. **Zikhali, Precious** (2008), Land Reform, Trust and Natural Resource Management in Africa
179. **Tengstam, Sven** (2008), Essays on Smallholder Diversification, Industry Location, Debt Relief, and Disability and Utility
180. **Boman, Anders** (2009), Geographic Labour Mobility – Causes and Consequences
181. **Qin, Ping** (2009), Risk, Relative Standing and Property Rights: Rural Household Decision-Making in China
182. **Wei, Jiegen** (2009), Essays in Climate Change and Forest Management
183. **Belu, Constantin** (2009), Essays on Efficiency Measurement and Corporate Social Responsibility
184. **Ahlerup, Pelle** (2009), Essays on Conflict, Institutions, and Ethnic Diversity
185. **Quiroga, Miguel** (2009), Microeconomic Policy for Development: Essays on Trade and Environment, Poverty and Education
186. **Zerfu, Daniel** (2010), Essays on Institutions and Economic Outcomes
187. **Wollbrant, Conny** (2010), Self-Control and Altruism
188. **Villegas Palacio, Clara** (2010), Formal and Informal Regulations: Enforcement and Compliance
189. **Maican, Florin** (2010), Essays in Industry Dynamics on Imperfectly Competitive Markets