The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Exit Dynamics of Start-up Firms: Does Profit Matter?

Rolf Golombek
Arvid Raknerud

CESifo Working Paper No. 5172
Category 12: Empirical and Theoretical Methods
January 2015

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5172*

Exit Dynamics of Start-up Firms: Does Profit Matter?

Abstract

We estimate by means of indirect inference a structural economic model where firms' exit and investment decisions are the solution to a discrete-continuous dynamic programming problem. In the model the exit probability depends on the current capital stock and a measure of short-run profitability, where the latter is a state variable which is unobserved to the econometrician. We estimate the model on all start-up firms in the Norwegian manufacturing sector during 1994-2012, and find that both increased short-run profitability and a higher capital stock lowers the exit probability - this effect is statistically significant in all industries. We show that the difference in annual exit probability between firms that exited during the observation period and firms that did not exit is highly persistent over time, and there is no tendency for a sharp increase in the estimated exit probability just prior to exit. Hence, it is the cumulated effect of higher risk of exit over several years - compared with the average firm - that causes exits.

JEL-Code: C330, C510, C610, C720, D210.

Keywords: exit, investments, indirect inference, continuous-discrete choice, monopolistic competition, costly reversibility.

Rolf Golombek
The Ragnar Frisch Center for Economic Research / University of Oslo
Oslo / Norway
rolf.golombek@frisch.uio.no

*Arvid Raknerud**
Statistics Norway
Research Department
Oslo / Norway
arvid.raknerud@ssb.no

*corresponding author

This paper has benefited from numerous comments and suggestions. In particular, we would like to thank Daniel Bergsvik, Erik Biørn, Bernt Bratsberg, John K. Dagsvik, Erik Fjærli, Torbjørn Hægeland, Jos van Ommeren, Knut Røed, Terje Skjerpen and Steinar Strøm. Earlier versions of the paper have been presented at the University of Oslo, at the Norwegian School of Management and at European Economic Association meeting in Toulouse 2014. We thank the participants for their comments. This research has been .nancially supported by The Norwegian Research Council (Grants no. 154710/510 and 183522/V10).

1 Introduction

Reallocation of resources from old, inefficient firms to new firms with superior technology is often considered to be the dynamo in a market economy; through creative destruction the exit of firms is a means to ensure growth and prosperity. New firms have to invest to build up an optimal stock of capital, but new firms are also characterized by a high exit rate: in our data set, which covers firms in Norwegian manufacturing industries over the period 1994-2012, the average probability that an one-year old firm exits during the next three years is 17 percent, compared to 7–8 percent for a 10-year old firm. For a rational firm, choosing the investment profile over time is interrelated with the decision of whether to exit today or continue production. Still, most theoretical as well as empirical studies solely examine either exit or investment. One contribution of the present paper is to derive a theory-based econometric model of firm exit and investment that is structurally estimated to obtain exit probabilities of firms.

In our dynamic model, the firm's investment decision is determined simultaneously with the decision of whether to exit. In contrast, almost all theoretical models of investment under uncertainty either rule out the possibility of exit or consider the value of exit – the "scrap value" – as exogenous. One important example is Dixit and Pindyck (1994; Chapter 7) who introduce the simplifying assumption that an investment project can be abandoned at a lump-sum cost, and also restarted at another lump-sum cost. Thus the firm can switch from one discrete state to another. Because none of these states are absorbing, the firm never really exits.

In the investment model in Bloom, Bond and van Reenen (2007), it is explicitly stated that exit is not an option. In Abel and Eberly (1994; 1996) a firm chooses positive, zero or negative investment according to the value of a state variable – the shadow price of capital. The firm may disinvest its entire stock of capital, but such an action does not lead to an absorbing state for the firm. Thus exit is de facto ruled out.

In another strand of the literature, exit and investment are considered simultaneously. Some prominent examples are Olley and Pakes (1996), whose method to estimate production functions is implemented in Stata and widely used, and Levinsohn and Petrin (2003). These authors specify models in which exit and investment are endogenous decisions, but if the firm exits it obtains a scrap value which is state independent, that is, *independent* of the firm's capital stock. In our model we replace this simplifying assumption by modeling a trade-off between the value of installed capital if production is continued and the value of installed capital if the firms exits – this is how we make the decision to exit truly endogenous.

While modeling of exit may seem simple – according to standard economic theory negative profitability is the key reason for firms to exit – our data indicate that exit behavior of firms may be more complicated: for the period 1994-2012 the data reveal that i) 27 percent of firms that exited had positive *profit* (here defined as operating surplus less capital costs) *in every year* before they exited, ii) there is no negative profitability shock just prior to exit; around 65 percent of the firms that exited had positive profit in the last year prior to exit, and iii) firms may continue production even though they repeatedly experience negative profit; 30 percent of the firm-year observations for the non-exiting firms – one observation for each firm in each year – had negative profit. These observations raise the following questions: Is profitability of key importance for explaining firm exit? What cause firms to exit? What are the characteristics that distinguish firms that exit from firms that continue production? Thus one purpose of the present paper is to identify, through estimating a dynamic structural microeconometric model, the answers to these questions.

Empirical papers on dynamic structural models of firms' investment or exit typically do not lead to numerically tractable criterion functions that can form the basis for estimation. Instead they often apply the simulated method of moments to estimate the structural parameters, see, for example, Cooper and Haltiwanger (2006), Hennessy and Whited (2007), Acemoglu et al. (2013) and Asphjell et al. (2014). Here the econometrician selects a set of moments ad hoc and let the parameters be determined such that the distance between the data moments and the corresponding model-based (simulated) moments are minimized according to some metric.

We are neither able to derive a likelihood function from our structural model that is numerically tractable. However, instead of using the simulated method of moments we introduce an auxiliary model that closely mimics the properties of the underlying structural, i.e., data-generating, model. In the auxiliary model, the probability to exit depends on a measure of (short-run) profitability and the stock of capital. The likelihood of the auxiliary model – the quasi-likelihood function – can be derived and therefore we estimate the parameters in the auxiliary model by maximum likelihood. The likelihood function of the auxiliary model – the quasi-likelihood function – can be derived and quasi-maximum likelihood estimates are combined with the structural model through simulations to estimate the parameters of the structural model.

The idea of combining estimation of an auxiliary model with simulations from an underlying "true" model is called indirect inference; see Gourieroux et al. (1993). One specific implementation of indirect inference is the efficient method of moments. In the present paper we draw on this approach, which was originally proposed by Gallant and Tauchen (1996). Indirect inference seems appropriate for our study because computing the exact likelihood is not feasible, whereas simulation of the model is fairly simple. Indirect inference is widely used in financial econometrics; some examples are stochastic volatility-, exchange rate-, asset price- and interest rate modeling, see, for example, Gallant and Long (1997), Andersen and Lund (1997), Andersen et al. (1999), Bansal et. al. (2007) and Raknerud and Skare (2012). However, indirect inference is not commonly used to estimate structural models of firm dynamics. We demonstrate that indirect inference is a viable approach also in this case.

We make three contributions to the literature. First, we present a novel theory-consistent econometric model that within the framework of stochastic dynamic programming determines both exit and investment. As noted above, in the literature the interrelationship between investment and exit has been neglected or it has been assumed that the value to exit is state independent, that is, independent of the firm's stock of capital. As demonstrated in the present study, this is hardly a suitable assumption.

Second, we contribute to the literature on the causal relationship between profitability and exit: we examine whether profitability is of key importance for explaining firm exit and we also identify the characteristics that distinguish firms that exit from firms that continue production. According to economic theory, exit is closely related to profitability, although the exact relationship varies between theories. The simplest theory suggests a myopic exit rule: "production is likely to come to a sharp stop ...[when] ... the price falls so low that is does not pay for the out of pocket expenses", Marshall (1966, p. 349). A more sophisticated theory suggests that the exit decision is based on both present and expected profits. The most refined theory derives the exit rule from stochastic dynamic programming: under the assumption that the firm takes into account that it will always make optimal decisions in the future, it will stay operative as long as the expected present value of continuing production exceeds the value to exit, see, for example, Hopenhayn (1992). We use a dynamic model that builds on stochastic dynamic programming, and show that it is the cumulated effect over several years of a high risk to exit that distinguishes firms that exit from firms that continue production; if this cumulated effect is sufficiently high, a firm exits.

Surprisingly, there is not much evidence in the literature on the relationship between profitability and exit. Some studies provide descriptive statistics on exit rates, see Dunne et al. (1988) for U.S. manufacturing industries and Disney et al. (2003) for UK manufacturing, but these studies do not provide information on the relationship between profitability and exit – the reason may be lack of data.

There is, however, a literature where reduced form probit models are used to examine how *profit components* have impact on firm exit. For example, Olley and Pakes (1996) analyze the evolution of plant-level productivity, but they also estimate how firm exit depends on age, the stock of capital and productivity. They find that a higher stock of capital, and also improved productivity, tend to decrease the exit probability. Foster et al. (2008) use a probit model and find that improved physical productivity, higher output prices and a higher stock of capital all tend to decrease the probability to exit.[1]

In contrast to Olley and Pakes (1996) and Foster et al. (2008), our analysis focuses on entrepreneurial firms. Thus new firms are included from the year they are born, while incumbent firms (at the start of the sample period) are excluded. The reason is that the exit probability of an incumbent firm may differ systematically from that of a new firm due to self-selection: the *surviving* firms are not a random sample of the population of all firms. In the literature, this selection problem is largely ignored. Estimates of the partial effect on firm exit of variables that are correlated with survival, such as productivity, age and size (capital stock), may therefore be biased.

Third, we shed new light on the role of size as an exit determinant. Typically, earlier

[1] The Foster et al. study is part of a growing literature where rich data bases are used to examine different aspects of firm productivity; two recent examples are Hsieh et al. (2009) and Bartelsman et al. (2013).

studies found that the probability to exit is higher the smaller the firm; some examples are Mata et al. (1995), Olley and Pakes (1996), Agarwal and Audretsch (2001), Klepper (2002), Disney et al. (2003), Pérez et al. (2004) and Foster et al. (2008). In our theory model that integrates investment and exit, a higher stock of capital has two opposite effects: More capital will increase production, and therefore raise the value of the firm if it continues to operate – this tends to lower the exit probability. On the other hand, more capital increases the scrap value of the firm, that is, the amount of money obtained if the firm sells its entire stock of capital – this tends to increase the exit probability. We show that with costly reversibility of investment, the first effect always dominates.

The rest of this paper is organized as follows: In Section 2, we identify stylized facts about the firms in the data set: these are firms in manufacturing industries (1994-2012). We show that adjustments of labor and materials from one year to the next exhibit a different pattern than adjustment of capital. Further, in all industries we observe huge aggregated profits over time. This suggests firms have market power, and we therefore assume imperfect competition (here modeled as monopolistic competition).

In Section 3 we introduce a production model – production requires input of labor, materials (including energy) and capital. The empirical observations in Section 2 justify to model materials and labor as fully flexible factors of production, whereas capital is assumed to be quasi-fixed with costly reversibility of investment; see Abel and Eberly (1996).

In Section 4 we explain how stochastic dynamic programming can be used to simultaneously determine (in each period) whether the firm will exit or not and how much the firm will invest if it does not exit. We extend Rust (1994) by allowing for, like in his original model, a discrete decision variable – whether or not to exit – in addition to a continuous decision variable – investment. We allow for both positive and negative investment. In particular, if a firm exits, it sells its entire stock of capital. Under the standard assumption that the state vector is Markovian, we derive the exit probability function of the firm. This is a function of the firm's scrap value – obtained if the firm exits – and the net present value of the firm if it continues production at least one more year and makes optimal decisions now and in the future.

We discuss the stochastic specification of the auxiliary econometric model in Section 5 and derive the quasi-likelihood function. The indirect inference estimator of the structural coefficients are presented in Section 6. The model estimates are reported in Section 7. We find that for a given level of capital, improved short-run profitability reduces the exit probability and this effect is statistically significant in all industries. The effect of a higher stock of capital (for a given level of short-run profitability) depends on two opposite effects, but our results confirm the theoretical prediction that the net effect is a lower exit probability. We also show that firms that exited during the observation period have a substantially higher estimated exit probability than firms that did not exit. The difference between estimated annual exit probabilities is highly persistent over time and is not limited to the year just prior to exit. In fact, the exit probabilities do not increase sharply just prior to exit, which reflects that there are no (negative) profitability shocks in the last years prior to exit. Therefore, it is the *cumulated effect of higher risk of exit* over several years – compared with the average firm – that causes exits. Finally, Section 8 concludes.

2 Data

Our main data source is a database from Statistics Norway based on register data – the Capital database – which covers the entire population of Norwegian limited liability companies in manufacturing. The main statistical unit in this database is the firm: A firm is defined as "the smallest legal unit comprising all economic activities engaged in by one and the same owner". We use data from the Capital database for the period 1993-2012.

We analyze the survival and dynamics of new firms as opposed to incumbent firms. A firm is defined to have entered in year $t - 1$ if it was first registered in the Capital data base in $t - 1$ and it was recorded also in year t. Further, a firm is defined to have exited in year t if it is recorded in the Capital database in year $t - 1$, but not in year t, *and* is registered as either bankrupt or having closed down for an unspecified reason after $t - 1$ according to the Central Register of Establishments and Enterprises (REE).[2] Note that a firm is removed from the Capital data base if it is no longer classified to belong to a manufacturing sector.

[2]There may be a delay in the registration of close downs in the REE – typically one or two years after the firm drops out from the Capital data base. This is the reason we have 2012 as our last data year.

We limit attention to new firms that were operative in at least two years. For each firm (that was operative at least two years), we use the first observation year solely to obtain information about the initial stock of capital of firms (at the end of that year).

We only include firms that are single-plant firm in the start-up year because newly established multi-plant firms are likely to be continuation of existing establishments under a new organization number (our firm identifier). In the period 2004-12, about 90 percent of the start-up manufacturing firms were single-plant units. These firms accounted for about two-third of total employment of all start-up firms in their first year. Finally, if a (single-plant) firm A acquires a (single-plant) firm B, then the new multi-plant firm A is kept in the data (whereas B is of course removed).

The Capital database contains annual observations on revenue, wage costs, intermediates expenses (including energy), fixed capital (tangible fixed assets) and many other variables for all Norwegian limited liability manufacturing firms for the period 1993-2012.[3] The database combines information from two sources: (i) accounts statistics for all Norwegian limited liability companies, and (ii) structural statistics for the manufacturing sector. In general, all costs and revenues are measured in nominal prices, and incorporate taxes and subsidies, except VAT. Labor costs include salaries and wages in cash and kind, social security and other costs incurred by the employer.

A unique feature of the database is that it contains the net capital stock in both current and fixed prices at the firm level. The data set distinguishes between two types of capital goods: (i) buildings and land, and (ii) other tangible fixed assets. The latter group consists of machinery, equipment, vehicles, movables, furniture, tools, etc., and is therefore quite heterogeneous. The method for calculating capital stocks in current prices is based on combining gross investment data and book values of the two categories of fixed tangible assets from the balance sheet, see Raknerud, Rønningen and Skjerpen (2007).

Our econometric model contains only a single aggregate capital variable. It has been constructed using a Törnqvist volume index, where each type of capital is proportional to the sum of: (i) the user cost of capital owned by the firm, and (ii) total leasing costs. This aggregation corresponds to a constant returns to scale Cobb-Douglas aggregation function for different types of capital (see OECD, 2001).[4]

[3]See Raknerud, Rønning and Skjerpen (2004).

[4]Formally, the aggregate capital stock is calculated using the Törnqvist volume index K_{it} =

Table 1 presents summary statistics for the five largest manufacturing industries and for the whole manufacturing sector when all firms are lumped together. The four industries we examine are Wood products (NACE 16), Metal products (NACE 25), Electrical equipment (NACE 27), Machinery (NACE 28), and Transport equipment (NACE 29-30). In the table the first and second column shows number of firms and number of exits by industry for the period 1994-2012. Column three depicts annual exit frequencies; these are typically 4-5 percent. The fourth column in Table 1 shows both the average and the median number of man-years in the entry year of firms. For total manufacturing, the mean is 14 and the median is 3. Among the individual industries, Transport equipment stands out with a high mean (38) and a median of 6 (man-years). Thus most firms are small – this is a typical feature of Norwegian manufacturing.

Firms in the manufacturing industries compete extensively at international markets. We therefore follow the standard in the international trade literature and assume imperfect competition, here specified as monopolistic competition. The basic idea of this assumption is that firms have some degree of market power, yet there are so many firms in the industry that it is reasonable to assume that each firm neglects that its choice of price has impact on the demand curve of its competitors.

Standard economic theory suggests that profit is (much) larger under imperfect competition - price exceeds marginal cost - than under perfect competition - price equal to marginal cost. As an informal test of our market structure assumption (monopolistic competition) we calculated wage costs, capital costs and profit aggregated over all firms in all periods (for each industry), and divided each of these by aggregated value added; the corresponding shares are shown in Table 1.[5] We find that profit make up between 8 and 12 percent of value added in the six industries.[6] Because perfect competition can

$(K_{it}^b)^{\nu}(K_{it}^o)^{(1-\nu)}$ where K_{it}^b and K_{it}^o are the stocks of buildings and land (b) and other tangible fixed assets (o). Further, $v = \sum_{it} R_{it}^b / \sum_{it}(R_{it}^b + R_{it}^o)$ where $R_{it}^k = (r + \delta_k)K_{it}^k$, $k = b, o$, is the annualized (user) cost of capital (including leased capital). In the latter expression r is the real rate of return, which we calculated from the average real return on 10-years government bonds for the period 1994-2009 (4 per cent), and δ_k is the median depreciation rate obtained from accounts statistics, see Raknerud, Rønningen and Skjerpen (2007). Because we have a single capital variable in the econometric model, we also have a single depreciation rate. This rate ($\delta = 12$ percent) is a weighted average of δ_k and δ_k with v as the weight.

[5] Capital costs are here calculated from the standard user cost formula with interest rate equal to the average yield on 10-years government bonds (see also footnote 4).

[6] According to the seminal paper by Mehra and Prescott (1985), risk aversion explains at most one percentage point of the US equity premium, that is, the difference between the return on equities and risk free bonds. This suggests that correcting for risk aversion will not alter the general picture in Table

be seen as a special case of the monopolistic competition model (infinitely large demand elasticity and a homogeneous good), in Section 7.1 we use our estimates to provide more evidence that perfect competition is not an adequate description of the market structure.

1.

Table 1: **Descriptive statistics for 1994-2012**

Industry (NACE)	No. of firms	No. of exits	Average exit-frequency*	Mean/median man-years**	Share of value added by: labor	capital	profit***
Wood products (16)	809	230	.048	11/3	.72	.20	.08
Metal products (25)	1246	296	.039	11/4	.74	.16	.10
Electrical equipment (27)	282	66	.038	16/3	.73	.12	.15
Machinery (28)	738	209	.045	12/3	.74	.14	.12
Transport equipment (29-30)	415	104	.043	38/6	.77	.13	.10
Total manufacturing	7419	2035	.043	14/3	.74	.16	.10

*Number of exits divided by number of firm-years

**Number of man-years at the year of entry

***Labor costs, (annualized) capital costs and profit as a share of value added

In Figure 1 we have examined how the use of labor (measured as man-hours), materials (intermediate inputs, including energy) and capital change over time. For each factor of production and each firm in each year, we first calculate the use of a factor in year t ($t = 1995, ..., 2012$) relative to the use of this factor in year $t - 1$. In Figure 1, the horizontal axis measures the log of this ratio, that is, the relative change in the use of inputs, and the vertical axis measures frequency. As seen from the figure, the graphs for man-hours and materials are almost identical and resemble the normal distribution. At first glance the graphs may give the impression that changes in man-hours and materials follow each other almost perfectly. There is, however, substitution possibilities between these two inputs: when comparing, for each industry, the within-firm variation in (log of) the materials-labor ratio to the within-firm variation in (log of) man-hours, we find that this ratio is around 50 percent. If materials and labor were used in a fixed ratio, specific to each firm, this ratio would have been 0. (This would also hold if the firm-specific ratios change proportionally over time for all firms). In Section 3 we therefore assume substitution possibilities between labor and materials.

Figure 1 also shows the graph for log of changes in the stock of capital. This graph has somewhat thicker tails than the graphs for man-hours and materials. The thicker tails mean that observations with large (negative or positive) changes are more frequent. Moreover, the thicker right tail – the graph is skewed to the right – reflects the intermittent and lumpy nature of investment in Norwegian manufacturing, see Nilsen and Schiantarelli (2003).

We see that net investment takes negative values for roughly 50 percent of the observations. A firm with negative net investment has lower acquisition of capital than depreciation. In particular, strongly negative net investment reflects sales of capital. In our data the value of annual sales of capital amount to around 10 percent of gross (annual) investment, which is substantial relative to aggregate depreciation. The distinct pattern of investment calls for another modeling of capital than of labor and materials, see Section 3.

In our data set a substantial share of the observations has negative profitability. This is the case both for i) firms that did not exit in the observation period ("non-exiting firms"), and ii) firms that did exit during the observation period ("exiting firms"). In fact, almost

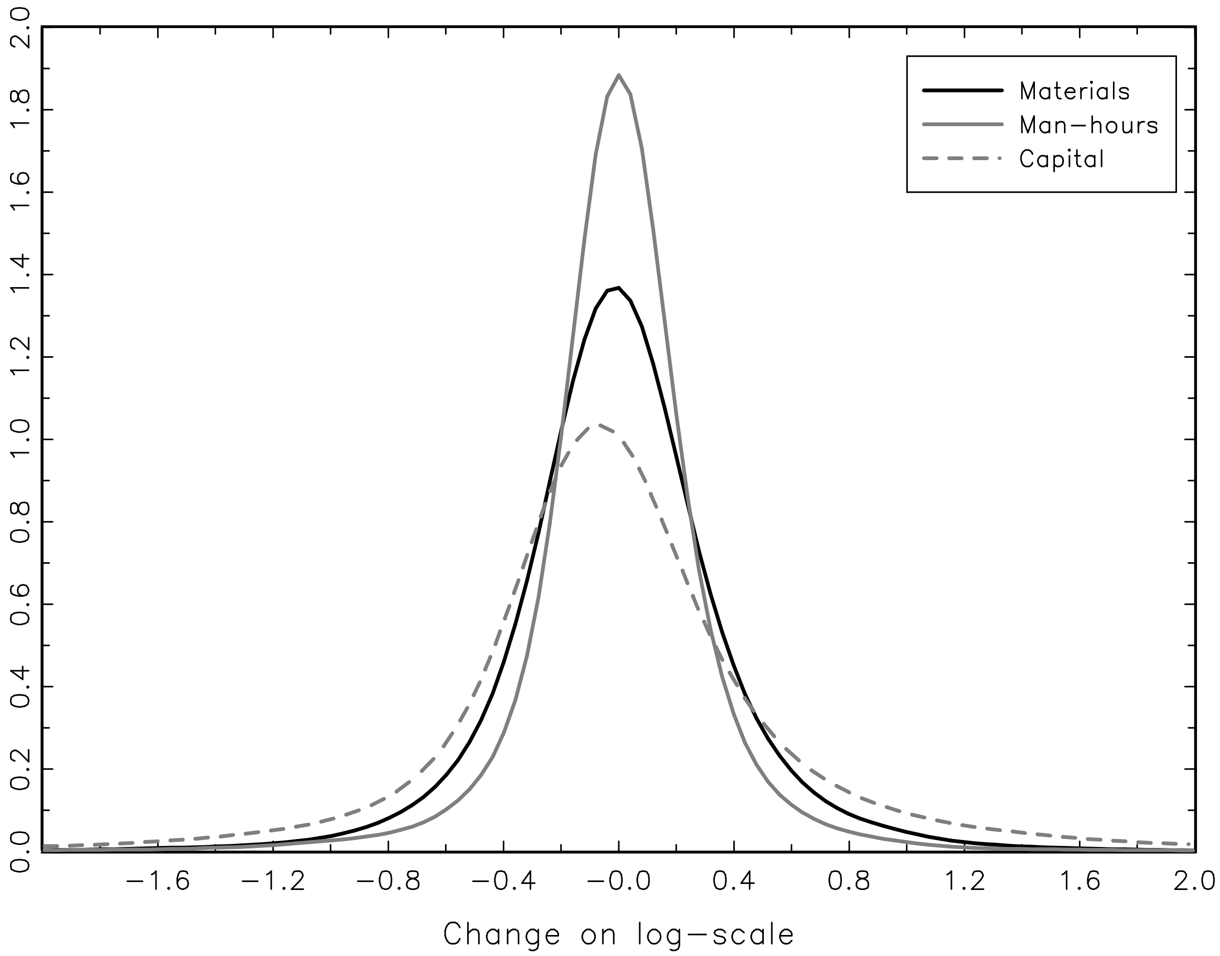

Figure 1: Distribution of log of annual changes in capital, man-hours and materials. Kernel density estimates. Total manufacturing, 1994-2012.

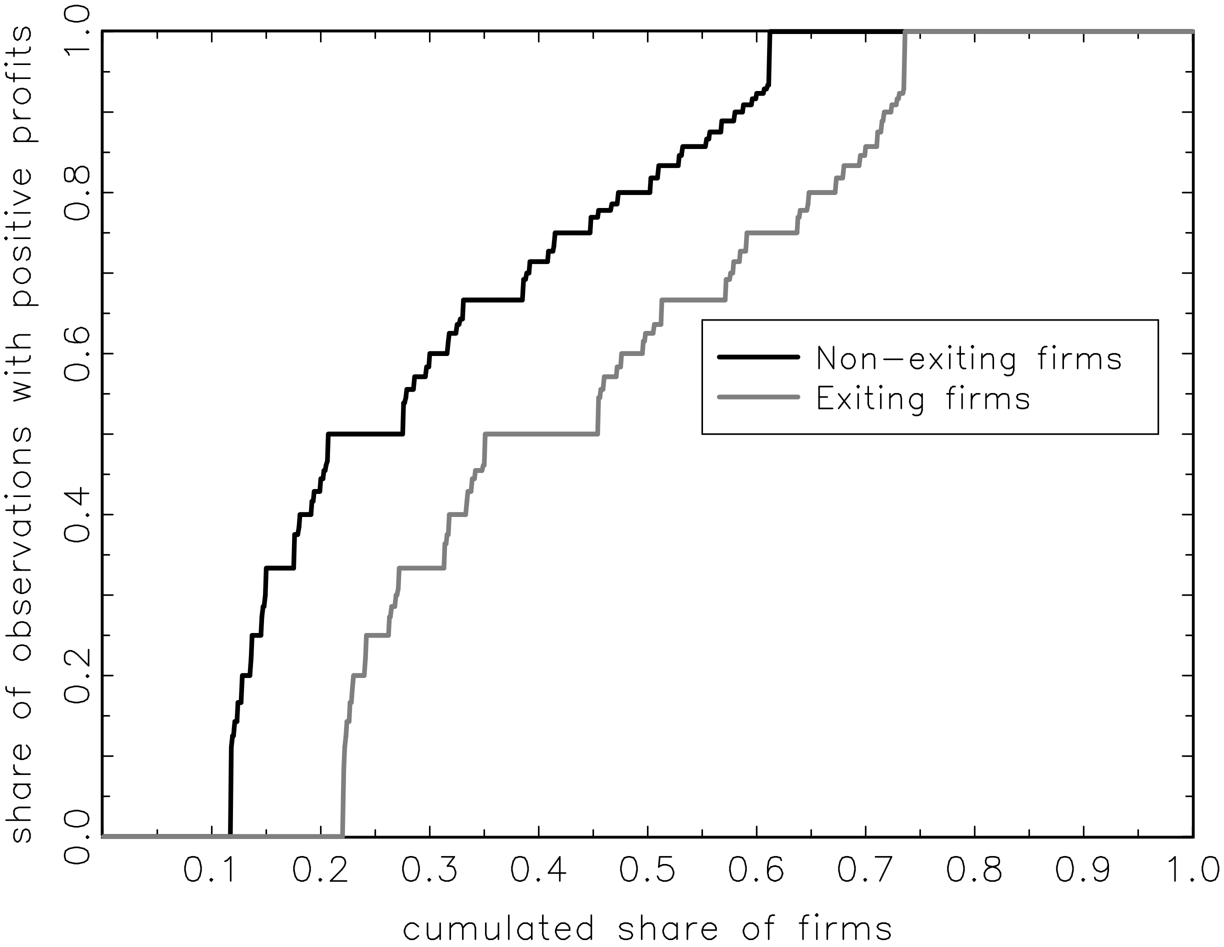

Figure 2: Distribution of share of observations (for each firm) with positive profits. Total manufacturing, 1994-2012.

20 percent of the firm-year observations of the non-exiting firms (one observation for each operating firm in each year), and more than 25 percent of the observations of the exiting firms, have negative operating surplus. The corresponding numbers for profit, that is, operating surplus less of capital costs, are 30 percent for non-exiting firms and 35 percent for exiting firms. Our model should therefore allow for negative profitability, in particular negative operating surplus.

The share of observations with negative profitability may be unevenly distributed over firms; some firms may have no, or just a few, observations with negative profitability, whereas others may have several observations with negative profitability. Figure 2 shows, for non-exiting and exiting firms, how the observations with *positive* profitability are distributed over firms. Each curve is constructed as follows: For each firm we find its share of observations with positive profitability, henceforth termed the positive profitability share. We then sort firms by their positive profitability share (from 0 to 1), and group firms with the same positive profitability share together. In Figure 2 the horizontal axis measures the cumulative share of firms, whereas the vertical axis measures the positive profitability share. Each curve consists of a number of steps. The length of each step shows the share of firms with the same positive profitability share, and the height of the step shows the positive profitability share.

Figure 2 shows that when measuring profitability by *profit*, about 22 per cent of the exiting firms have a positive profitability share of zero, that is, all their observations have negative profit. The corresponding number for the non-exiting firms is 12 percent. Almost 45 percent of the exiting firms have a positive profitability share that is 0.5 or lower, i.e., at least half of their observations have negative profit. We also see that around 27 (38) percent of the exiting (non-exiting) firms have a positive profitability share of 1, that is, they have positive profit in every year.

Figure 2 gives a mixed picture of the importance of profitability relative to exit. On the one hand, a substantial share of the exiting firms (27 percent) always have positive profit. Moreover, most exiting firms are profitable the last years before the exit: The share of exiting firms with positive operating surplus three years, two years and one year prior to exit was 86 percent, 82 percent and 75 percent, respectively. The corresponding shares with positive profits are about 10 percentage points lower. On the other hand, the graph of the non-exiting firms lies above that of the exiting firms, reflecting that the former group on average has the highest profitability. The area between the two graphs is considerable, suggesting that there is a negative relationship between profitability and exit. We return to the question of whether there is a significant relationship between profitability and exit in Section 7.2.

3 Short-run factor demand

In this section we present our model for price decisions by firms: We consider an industry with monopolistic competition. Each producer faces a demand function of the following form:

$$Q_{it} = \Phi_t P_{it}^{-e} \tag{1}$$

where Q_{it} is output from firm i at time t, P_{it} is the output price and Φ_t is an exogenous demand-shift parameter characterizing the size of the market. Furthermore, $e > 1$ is the absolute value of the direct price elasticity. The price elasticity is common to all firms and constant over time.

Let M_{it} denote materials, L_{it} labor, and K_{it} capital. In Section 2 we argued that the modeling of materials and labor should be similar, but this modeling should differ from the one for capital. We now assume that the use of materials and labor are determined at the beginning of a time period (variable inputs), whereas capital services in year t are determined by the capital stock at the end of $t-1$; $K_{i,t-1}$. However, through investment in period t the capital stock at the end of period t increases (capital is quasi-fixed – see discussion below). The production function of producer i is assumed to be:

$$Q_{it} = A_{it} K_{i,t-1}^{\gamma} \left[M_{it}^{\rho} + (w_t L_{it})^{\rho} \right]^{\frac{\varepsilon}{\rho}}, \; \rho < 1 \tag{2}$$

where the elasticity of scale is equal to $\varepsilon + \gamma$, the elasticity of substitution between materials and labor is $1/(1-\rho)$ and w_t is a time-varying distribution parameter. Our production function is a nested Cobb-Douglas function defined over capital and a CES aggregate over labor and materials. The specification (2) allows for heterogeneity in productivity across firms: Hicks-neutral changes in efficiency are picked up by A_{it}, which may shift over time and vary across firms, whereas a positive change in w_t can be interpreted as a labor-augmenting innovation. Thus w_t captures that the skill-composition of labor typically changes over time. Whereas L_{it} is the use of labor measured in man hours, $w_t L_{it}$ should be interpreted as the use of labor measured in efficiency units.

Let $\mathbf{q}_{it} = (q_{Mt}, q_{Lit})$ be a vector of the unit price of materials and labor, respectively. The unit price of labor is firm specific, which reflects that the composition of different types of labor may vary across firms. All prices have been deflated and are thus real prices. We deflate all prices by the same index, so that in any time period one dollar of any cost component has the same value as one dollar of a revenue component. (If profit components are deflated by different indexes, nominal profit and deflated profit may have different signs.) We use the price index of capital, q_{Kt}, as the deflator to reflect the opportunity cost of investment.

Producers are assumed to be price takers in all factor markets. Using Shephard's lemma, the short-run cost function can be shown to be

$$C(\mathbf{q}_{it}, K_{i,t-1}, Q_{it}) = c_{it} \left(\frac{Q_{it}}{A_{it} K_{i,t-1}^{\gamma}} \right)^{\frac{1}{\varepsilon}} \tag{3}$$

where

$$c_{it} = [q_{Mt}^{\varrho} + (q_{Lit}/w_t)^{\varrho}]^{\frac{1}{\varrho}} \, , \; \varrho = \frac{\rho}{\rho - 1}. \tag{4}$$

Here, c_{it} is a firm-specific price index of variable inputs, i.e., it is derived from the CES-aggregate of materials and labor. Note that c_{it} depends on the distribution parameter w_t; q_{Lit}/w_t is the efficiency corrected price of labor.

The short-run optimization problem of firm i in the beginning of period t, when the producer knows $\mathbf{q}_{it}$, Φ_t, A_{it} and w_t (and also e, γ, ρ and ε), is to choose - for a given stock of capital - the price that maximizes operating surplus:

$$\Pi_{it} = \max_{P_{it}} \left\{ \Phi_t P_{it}^{1-e} - c_{it} \left(\frac{\Phi_t P_{it}^{-e}}{A_{it} K_{i,t-1}^{\gamma}} \right)^{\frac{1}{\varepsilon}} \right\} \tag{5}$$

where $\Phi_t P_{it}^{1-e} = P_{it} Q_{it}$ (from (1)) is the revenue of the firm. Solving the resulting first-order condition gives the following equations for revenue $R_{it} = P_{it} Q_{it}$ and short-run factor costs $q_{Mt} M_{it}$ and $q_{Lit} L_{it}$:

$$\begin{aligned} \begin{bmatrix} \ln R_{it} \\ \ln(q_{Mt} M_{it}) \\ \ln(q_{Lit} L_{it}) \end{bmatrix} &= \begin{bmatrix} \theta_2 \\ \theta_2 - \varrho \\ \theta_2 - \varrho \end{bmatrix} \ln c_{it} + \begin{bmatrix} 0 \\ 0 \\ \varrho \end{bmatrix} \ln(q_{Lit}/w_t) + \mathbf{1}\theta_1 \ln A_{it} \\ &\quad + \mathbf{1}\gamma\theta_1 \ln K_{i,t-1} + \varrho \begin{bmatrix} 0 \\ \ln q_{Mt} \\ 0 \end{bmatrix} + \mathbf{1} d_t, \end{aligned} \tag{6}$$

where $\mathbf{1}$ is a vector of ones,

$$d_t = \frac{\theta_1}{e-1} \ln \Phi_t \,.$$

and

$$\theta_1 = \frac{(e-1)}{(\varepsilon + e - e\varepsilon)} > 0,\ \ \theta_2 = \frac{-\varepsilon(e-1)}{(\varepsilon + e - e\varepsilon)} < 0. \tag{7}$$

We see that θ_1 is the coefficient of the Hicks-neutral efficiency term $\ln A_{it}$, which is common in all the three equations in (6). In contrast, a change in the firm-specific price index of variable inputs, c_{it}, will have a different impact on revenues (θ_2) than on factor costs ($\theta_2 - \varrho$). Note that an increase in w_t (for given q_{Lit}) increases revenue R_{it} because $\theta_2 < 0$, see (6) and (7). An increase in w_t has no direct impact on material costs, see (6), but will, through a drop in the firm-specific price index c_{it} (see (4)), increase material costs if $\theta_2 < \varrho$, see (6). An increase in w_t has an identical indirect effect, through c_{it}, on material costs as on labor costs ($\theta_2 - \varrho$), but has in addition a direct impact on labor costs (ϱ). If $\varrho > 0$, an increase in w_t will therefore lower the short-run cost share of labor, i.e., the innovation is labor saving.

Note that if the demand parameter is allowed to be firm-time specific, denoted Φ_{it}, the system (6) is *unaltered* except that A_{it} is replaced by $A^*_{it} = \Phi_{it}^{1/(e-1)} A_{it}$. Thus, neutral efficiency shocks (A_{it}) and (idiosyncratic) demand shocks (Φ_{it}) enter the two alternative systems in a completely symmetric way, and we would not be able to distinguish between them in the empirical analysis. Therefore, A_{it} captures both technology shocks and demand shocks, but we will still refer to A_{it} as "efficiency". This should be kept in mind when interpreting the results reported in Section 7.

Operating surplus Π_{it} defined in (5) has the closed form

$$\begin{aligned} \Pi_{it} &= (1 - e^{\varrho(\ln q_{Mt} - \ln(c_{it}))} - e^{\varrho(\ln(q_{Lit}/w_t) - \ln(c_{it}))}) e^{d_t} c_{it}^{\theta_2} A_{it}^{\theta_1} K_{i,t-1}^{\gamma\theta_1} \\ &\equiv \pi_{it} K_{i,t-1}^{\gamma\theta_1}, \end{aligned} \tag{8}$$

where

$$\ln \pi_{it} = b_{it} + \theta_2 \ln c_{it} + d_{1t} + \theta_1 \ln A_{it} \tag{9}$$

and

$$b_{it} = \ln(1 - e^{\varrho(\ln q_{Mt} - \ln(c_{it}))} - e^{\varrho(\ln(q_{Lit}/w_t) - \ln(c_{it}))}). \tag{10}$$

In order to ensure that the optimization with respect to capital is well-defined, we need to have $\gamma\theta_1 < 1$. (Our model meets this requirement, see below). In Section 5 we describe in detail how we identify Π_{it} using data on revenue, variable factor costs and capital. From the implicit definition of π_{it} in (8) we see that this variable depends on a number of factors that have impact on the profitability of a firm. Below we will therefore refer to π_{it} as a measure of profitability. Finally, note that π_{it} does not depend on the stock of capital, see (8); it is a measure of short-run profitability.

4 Exit and investment dynamics

The producer invests in capital during year t. We follow the standard assumption that it takes one period until the stock of capital is adjusted. If there were no costs of adjusting capital, then the stock of capital is found from maximizing

$$\Pi_{it} - (r + \delta)K_{i,t-1} \tag{11}$$

with respect to $K_{i,t-1}$ where Π_{it} is a function of $K_{i,t-1}$ given by (8) and $(r + \delta)K_{i,t-1}$ is the (neoclassical) user cost of capital (r is the real interest rate, δ the depreciation rate and the price of capital is normalized to one because capital is the numeraire good, see discussion above). In this paper we will, however, allow for capital adjustment costs.

At the beginning of each year t, the firm makes an investment decision. Investment I_t can be positive or negative. In particular, if the firm decides to exit during year t, it will sell its remaining stock of capital at the end of year t; $I_t = -(1 - \delta)K_{t-1}$.

Let z_t be a dummy variable which is one if the firm continues to operate throughout year t and zero if the firm exits during year t. We take the Markovian discrete choice model of Rust (1994) as a starting point and assume that the period t utility from the choice (I_t, z_t), given the state vector $S_t = (\pi_t, K_{t-1})$, can be written as:

$$u(S_t, I_t, z_t) + \varepsilon(z_t) \tag{12}$$

where $u(S_t, I_t, z_t)$ is operating surplus minus capital expenditures and $\varepsilon(z_t)$ is a random component associated with the discrete choice z_t. By definition we have

$$u(S_t, I_t, z_t) = \begin{cases} \Pi_t - c(I_t) & z_t = 1 \\ \Pi_t - c(-(1 - \delta)K_{t-1}) & z_t = 0 \end{cases} \tag{13}$$

where the function $c(I_t)$ denotes total cost of capital. Below we will assume that there is one type of capital adjustment costs, namely that the resale price of capital is lower than the purchaser price of capital, i.e., costly reversibility (see Abel and Eberly, 1996).[7] Then $c(I_t)$ is weakly convex with a kink at zero. Operating surplus Π_t follows from S_t and is therefore not affected by z_t and I_t. If $z_t = 0$, t is the terminal period and the firm sells its remaining capital stock, $I_t = -(1-\delta)K_{t-1}$, and obtains a scrap value, $-c(-(1-\delta)K_{t-1})$, at the end of the year.

Following Rust (1994) we assume that the state vector S_t is Markovian with transition probability $g(dS_{t+1}|S_t, I_t)$ and that $\varepsilon(z) = (\varepsilon(0), \varepsilon(1))$ has a bivariate extreme value distribution with scale parameter τ and location parameters $\gamma_z = (\gamma_0, \gamma_1)$:[8]

$$h(\varepsilon) = \prod_{z\in\{0,1\}} \tau \exp\{-\tau\varepsilon(z) + \gamma_z)\} \exp\{-\exp\{-\tau\varepsilon(z) + \gamma_z\}\}. \tag{14}$$

Further, the firm's choice of whether to continue production, and if so, how much to invest, follows from the solution of the Bellman equation:

$$V(S_t, \varepsilon_t) = \max_{z_t,\, I_t} \left\{ u(S_t, I_t, z_t) + \varepsilon(z_t) + \frac{1}{1+r} E_t \left[V(S_{t+1}, \varepsilon_{t+1}) \right] \right\}. \tag{15}$$

The value function $V(S_t, \varepsilon_t)$ is characterized in Proposition 1, which is an extension of the discrete choice model of Rust (1994), that is, we allow for a discrete *and* a continuous decision variable.

Proposition 1 *Assume (12)-(14) and that S_t is Markovian with transition probability $g(dS_{t+1}|S_t, I_t)$. Then the expected net present value of the firm is*

$$V(S_t, \varepsilon_t) = \max_{z_t\in\{0,1\}} \left[\Pi_t + v(S_t, z_t) + \varepsilon(z_t) \right] \tag{16}$$

where

$$v(S_t, 0) = -c(-(1-\delta)K_{t-1}) \tag{17}$$

and

$$\begin{aligned} v(S_t, 1) \;\; = \;\; & \max_{I_t} \left\{ -c(I_t) + \frac{1}{1+r} \times \right. \\ & \left. \int \left[\Pi_{t+1} + \frac{1}{\tau} \ln \left[\exp(-\tau c(-(1-\delta)K_t) + \gamma_0) + \exp(\tau v(S_{t+1}, 1) + \gamma_1) \right] \right] g(dS_{t+1}|S_t, I_t) \right\}. \end{aligned} \tag{18}$$

[7] An alternative assumption is that total cost of capital also includes resources to adjust to a higher stock of capital. Under the standard assumption that this type of cost of adjusment is decreasing in the initial stock of capital (for a given level of investment), see Abel and Eberly (1994), all our results go through.

[8] Because $E(\tau\varepsilon(z) - \gamma_z) = \gamma$ where γ is Eulers' constant, we have $E(\varepsilon(z)) = (\gamma + \gamma_z)/\tau$.

Finally, the exit probability is given by

$$\Pr(z_t = 0|S_t, z_{t-1} = 1) = \frac{1}{1 + \exp\{-(-\tau[v(S_t,1) - v(S_t,0)] + \nu)\}}, \tag{19}$$

where $\nu = \gamma_0 - \gamma_1$.

The proof of Proposition 1 is given in the Appendix. In Proposition 1, $v(S_t, 1)$ is the net present value of the firm if it does not exit in the current period ($z_t = 1$) and makes optimal investment decisions now (I_t) and in the future:

$$v(S_t, 1) = \max_{I_t}\left\{-c(I_t) + \frac{1}{1+r}E_t\left[V(S_{t+1}, \varepsilon_{t+1})\right]\right\}.$$

Above we assumed that the resale price of capital is lower than the purchaser price of capital. This assumption is now specified as

$$c(I) = \begin{cases} I & \text{if } I \geq 0 \\ sI & \text{if } I < 0 \end{cases} \quad s \leq 1. \tag{20}$$

According to (20), upon selling capital ($I < 0$) the firm may not obtain the purchaser price of capital: Markets for old capital may be imperfect, or there may be large transaction costs, that is, $s < 1$. For parts of the capital stock there may even be no market (i.e., zero price) because of, for example, asymmetric information. In that case the firm will face clean-up costs when the old capital is removed from the production site. The assumption that $s < 1$ may be particularly relevant for a small country like Norway because of thin second-hand markets for capital. The special case $s = 1$ corresponds to the neoclassical theory of investment.

Let $S_t = (\pi_t, K_{t-1})$ and $S'_t, = (\pi_t, K'_{t-1})$ with $K'_{t-1} > K_{t-1}$. Then

$$v(S'_t, 1) \geq s(1-\delta)(K'_{t-1} - K_{t-1}) + v(S_t, 1).$$

We conclude that

$$\frac{v(S'_t, 1) - v(S_t, 1)}{K'_{t-1} - K_{t-1}} \geq s(1-\delta).$$

Because $v(S_t, 0) = s(1-\delta)K_{t-1}$, we must have

$$\partial(v(S_t, 1) - v(S_t, 0))/\partial K_{t-1} \geq 0, \tag{21}$$

implying that $v(S_t, 1) - v(S_t, 0)$ is non-decreasing in the current stock of capital. This suggests that the probability to exit is lower the higher the stock of capital. It is easy to

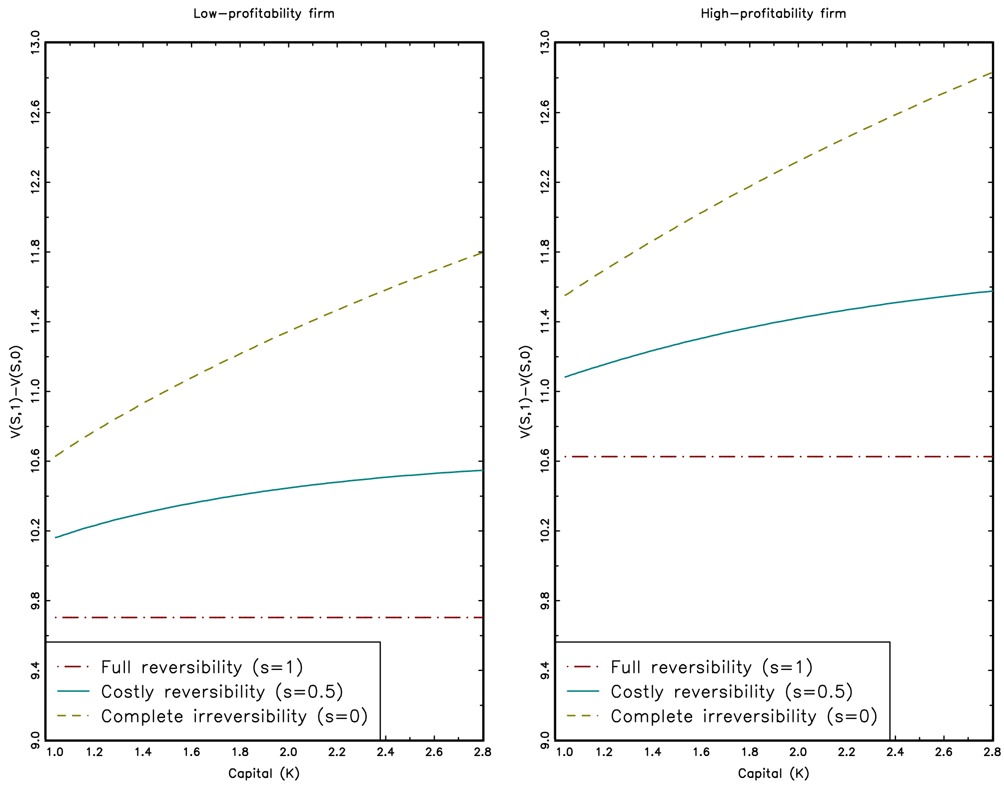

Figure 3: The net value of continuing as functions of capital (K_{t-1}) for three levels of adjustment costs (s) and two leveks of short-run profitability (π_t).

show that if $g(dS_{t+1}|S_t', I_t)$ stochastically dominates $g(dS_{t+1}|S_t, I_t)$ for all $S_t = (\pi_t, K_{t-1})$ and $S_t' = (\pi_t', K_{t-1})$ with $\pi_t' > \pi_t$,[9] then $\partial v(S_t, 1)/\partial \pi_t \geq 0$.

Figure 3 illustrates typical solutions of the value functions $V(S_t, 0)$ and $V(S_t, 1)$ and depicts the difference $V(S_t, 1) - V(S_t, 0)$ (the net value of continuing production) as a function of K_{t-1} for different values of s ($s = 0, .5, 1$) and π_t ("low profitability" and "high profitability"). In particular, we see that when $s = 1$ (no adjustment costs/full reversibility), $v(S_t, 1) - v(S_t, 0)$ does not depend on K_{t-1}. Furthermore, we see that when $s < 1$, $v(S_t, 1) - v(S_t, 0)$ is increasing in K_{t-1} for a given level of short-run profitability (π_t).

[9]That is, $G(S_{t+1}|S_t') \leq G(S_{t+1}|S_t)$ for any S_{t+1}, where $G(S_{t+1}|S_t)$ is the c.d.f. corresponding to the p.d.f. $g(S_{t+1}|S_t)$. In our model this means that a higher current profitability, π_t ,uniformly shifts the cumulative distribution function of next year's profitaility, π_{t+1}, rightwards.

5 Quasi-likelihood estimation

Our estimation strategy consists of two steps. In the first step, we specify an auxiliary model that approximates our structural model. The auxiliary model forms the basis for estimating the structural parameters by indirect inference. We denote the likelihood function of the auxiliary model for the quasi-likelihood function. The maximizer of the parameters, say ψ, of this quasi-likelihood function is the quasi-likelihood estimator, $\widehat{\psi}$.

In the second step, the parameters of the structural model, say θ, is estimated by simulating from the underlying "true" (data-generating) model. Our indirect inference estimator draws on the efficient method of moments estimator, see Gallant and Tauchen (1996). The estimator finds, through simulations of the economic model for a given θ, the value of θ that minimizes (in a weighted mean squared error sense) the score vector of the quasi-likelihood function for the simulated data when this score vector is evaluated at the quasi-likelihood estimator, $\widehat{\psi}$, obtained from the real data.

Measurement and identification issues Whereas the solution to (6) corresponds to an ex ante production plan that is based on the information available to the firm at the beginning of t, the ex post realizations, i.e., the data, are also determined by other (unpredictable) factors, for example, measurement errors and new information obtained during the year. In practice, the *observed* variables corresponding to the vector of theoretical variables will not satisfy the strong restrictions imposed by (6). Therefore, we will incorporate (non-structural) error terms into our model. Let

$$\mathbf{y}_{it} = \left(\ln \widehat{R}_{it}, \ln(q_{Mt}\widehat{M}_{it}), \ln(q_{Lit}\widehat{L}_{it})\right)'.$$

We assume that $\mathbf{y}_{it}$ is equal to the corresponding structural variables except for additive white noise error terms. That is

$$\mathbf{y}_{it} = \begin{bmatrix} \ln R_{it} \\ \ln(q_{Mt}M_{it}) \\ \ln(q_{Lit}L_{it}) \end{bmatrix} + \begin{bmatrix} e_{Rit} \\ e_{Mit} \\ e_{Lit} \end{bmatrix}. \tag{22}$$

In our data we observe firm specific wages, q_{Lit}, but only a price index for material costs q_{Mt}, which is normalized to one in the base year. Note that this is not a problem for our model. To see this, define $q^*_{Mt} = \lambda q_{Mt}$ for an arbitrary normalizing constant λ. Then define $w_t^* = w_t/\lambda$, $d_t^* = (\theta_1/e - 1)\ln\Phi_t - \theta_2 \ln\lambda$, and

$$c_{it}^* = [(q_{Lit}/w_t^*)^{\varrho} + q_{Mt}^{*\;\varrho}]^{\frac{1}{\varrho}}\,.$$

It is easy to show that (6) still holds with $(q_{Mt}, w_t, d_t, c_{it})$ replaced by $(q_{Mt}^*, w_t^*, d_t^*, c_{it}^*)$. Thus (6) is valid for any normalization of q_{Mt}.

Because A_{it} is unobserved, we cannot identify θ_1: Define $a_{it} = \ln A_{it}/\widetilde{k}$ for an arbitrary proportionality factor $\widetilde{k}$ and let $\widetilde{\theta}_1 = \widetilde{k}\theta_1$. Then

$$\widetilde{\theta}_1 a_{it} = \theta_1 \ln A_{it} \tag{23}$$

regardless of $\widetilde{k}$. The parameter $\widetilde{\theta}_1$ can be identified only by making stochastic assumptions about a_{it}. To obtain identification we assume that

$$a_{it} = \varphi a_{i,t-1} + \eta_{it},\ t = 2, ..., \tau_i \tag{24}$$

$$a_{i1} \sim \mathcal{IN}(0, \sigma_a^2),\ \eta_{it} \sim \mathcal{IN}\,(0, 1)\,. \tag{25}$$

These assumptions enable us to identify the loading coefficient $\widetilde{\theta}_1$, but tells us nothing about the structural parameter θ_1 since $\widetilde{k}$ is unidentified. By a similar argument, any non-zero mean in a_{it} would be absorbed into the term d_t, hence the assumption that a_{it} has zero mean is also a purely identifying restriction.

The variable a_{i1} represents the productivity of firm i in its start-up year relative to the average productivity of all new firms in that year, and the variance σ_a^2 of a_{i1} characterizes the cross-sectional heterogeneity across firms in their first observation year. Observed productivity differences among operative firms in a later year is the result of initial heterogeneity, a_i, cumulated innovations, $\sum_{t=2} \eta_{it}$, and self-selection (the most productive firms survive). In order to obtain identification, both the initial value of a_{i1} and the subsequent innovations η_{it} must have zero mean since any non-zero mean will be indistinguishable from the industry-wide intercept d_t in (9). Moreover, the variance of the innovation η_{it} is set to one to obtain identification of $\widetilde{\theta}_1$.

Below we show how to consistently estimate $\theta_2 = -\varepsilon(e-1)/(\varepsilon + e - e\varepsilon)$, see (7), by indirect inference. Still we we cannot identify both ε and e. To obtain identification of both ε and e we need to impose an additional condition. For example, if markets are assumed to be competitive, that is, $e \to \infty$, then $\theta_2 = -\varepsilon/(1-\varepsilon)$ and $\gamma\theta_1 = \gamma/(1-\varepsilon)$, so both ε and γ are identified. Alternatively, we can assume that the elasticity of scale is $\varepsilon + \gamma = 1$. Then $\theta_2/\gamma\theta_1 = -\varepsilon/(1-\varepsilon)$, so ε is identified and then e follows from θ_2.

An auxiliary model of exit There are two problems using the exit probability (19) to estimate the structural parameters. First, to solve the functional fixed point equation (18) for given S_t. This problem is difficult, but tractable. Second, S_t contains a latent state variable, π_t. To handle the second problem we approximate the structural model by an auxiliary model, that is, we approximate $v(S_t, 1) - v(S_t, 0)$ by a parametric function of observable variables:

$$v(\widehat{S}_t, 1) - v(\widehat{S}_t, 0) \simeq \beta_0^* + \beta_K^* K_{i,t-1}^{\gamma_K} + \beta_\pi^* \widehat{\pi}_{i,t-1}^{\gamma_\pi}. \qquad (26)$$

Here $\widehat{S}_{it} = (\widehat{\pi}_{i,t-1}, K_{i,t-1})$ is the observable equivalent of the state vector S_{it}, and

$$\widehat{\pi}_{it} = \max(\widehat{\Pi}_{it}/K_{i,t-1}^{\gamma\theta_1}, 0)$$

with $\widehat{\Pi}_{it}$ being observed operating surplus in the data year t. Because $\widehat{\Pi}_{it}$ is not defined when $z_{it} = 0$, we use $\widehat{\pi}_{i,t-1}$ instead of $\widehat{\pi}_{it}$ in (26). This can be justified if $\ln \pi_{it}$ is a Gaussian AR(1) process, see below. Then $E(\pi_{it}|\pi_{i,t-1})$ will be a power function in $\pi_{i,t-1}$, which motivates the use of a power function in (26). The restriction $\widehat{\pi}_{it} \geq 0$ is imposed because π_{it} cannot be negative; this is a well-known property of the (nested) Cobb-Douglas production function.

Combining (20) and (26) we can rewrite (19) as

$$\begin{aligned} \Pr(z_{it} &= 0|\ S_{it}, z_{i,t-1} = 1) \simeq \frac{1}{1 + \exp\left\{-\left[\beta_0 + \beta_K K_{i,t-1}^{\gamma_K} + \beta_\pi \widehat{\pi}_{i,t-1}^{\gamma_\pi}\right]\right\}} && (27) \\ &\equiv P(z_{it} = 0|\ K_{i,t-1}, \widehat{\pi}_{i,t-1}), && (28) \end{aligned}$$

where $\beta_0 = -\tau\beta_0^* + \nu$, $\beta_K = -\tau\beta_K^*$ and $\beta_{2\pi} = -\tau\beta_{2\pi}^*$. In Section 4 we derived that $\partial\left(v(S_t, 1) - v(S_t, 0)\right)/\partial K_{t-1} \geq 0$ (and independent of K_{t-1} if $s = 1$) and $\partial v(S_t, 1)/\partial \pi_t \geq 0$. This suggests that $\beta_K^* > 0$ and $\beta_\pi^* > 0$, and hence $\beta_K < 0$ and $\beta_\pi < 0$; both a higher capital stock and improved profitability lower the probability to exit. Note that if there are no adjustment costs of capital ($s = 1$), then $\beta_K = 0$.

Initial estimation We now consider the estimation of ϱ and w_t. From (6) we have

$$\ln\left(\frac{q_{Lit}L_{it}}{q_{Mt}M_{it}}\right) = -\varrho\ln w_t + \varrho\ln\left(\frac{q_{Lit}}{q_{Mt}}\right) + e_{Lit} - e_{Mit}. \tag{29}$$

We can utilize (29) to get simple regression estimates of ϱ and w_t; $(\widehat{\varrho},\ \widehat{w}_t)$. Hence, c_{it} can be estimated as:

$$\widehat{c}_{it} = \left[(q_{Lit}/\widehat{w}_t)^{\widehat{\varrho}} + q_{Mt}^{\widehat{\varrho}}\right]^{\frac{1}{\widehat{\varrho}}} \tag{30}$$

and b_{it} as

$$\widehat{b}_{it} = \ln(1 - e^{\widehat{\varrho}(\ln q_{Mt} - \ln(\widehat{c}_{it}))} - e^{\widehat{\varrho}(\ln(q_{Lit}/\widehat{w}_t) - \ln(\widehat{c}_{it}))}). \tag{31}$$

Henceforth, in all expressions where c_{it} and b_{it} enter we will replace them by $\widehat{c}_{it}$ and $\widehat{b}_{it}$, respectively, ignoring the approximation error.

Quasi-likelihood estimation Given the estimates obtained in the initial estimation and the reduced form structural model derived from the short-run factor adjustment, we can write:

$$\mathbf{y}_{it} = \begin{bmatrix} \widetilde{\theta}_1 \\ \widetilde{\theta}_1 \\ \widetilde{\theta}_1 \end{bmatrix} a_{it} + \begin{bmatrix} \theta_2 \\ \theta_2 - \varrho \\ \theta_2 - \varrho \end{bmatrix} \ln\widehat{c}_{it} + \begin{bmatrix} 0 \\ \varrho\ln q_{Mt} \\ \varrho\ln(q_{Lit}/\widehat{w}_t) \end{bmatrix}$$
$$+ \begin{bmatrix} \gamma\theta_1 \\ \gamma\theta_1 \\ \gamma\theta_1 \end{bmatrix} \ln K_{i,t-1} + \begin{bmatrix} d_t \\ d_t \\ d_t \end{bmatrix} + \mathbf{e}_{it} \text{ for } t = 1, ..., \tau_i. \tag{32}$$

where

$$\mathbf{e}_{it} \sim \mathcal{IN}(\mathbf{0},\Sigma_e). \tag{33}$$

In general, let the parameters of the quasi-likelihood be denoted ψ. In our model,

$$\psi = \left(\widetilde{\theta}_1, \theta_2, \gamma\theta_1, \sigma_1^2, \sigma_a^2, d_1, ..., d_T, vech(\Sigma_e)\right).$$

Given the estimates $\widehat{\varrho}$ and $\widehat{w}_t$ obtained in the first step of the estimation, our data on firm i can be seen as the realization of a stochastic process $(\mathbf{y}_{i1}, ..., \mathbf{y}_{i\tau_i})$ and $\tau_i \leq \overline{T}_i \leq 2012$ is the stopping time. Here, $\overline{T}_i$ is the firm-specific year of right censoring, which is exogenous. For simplicity of notation we have assumed that the firm enters at $t = 1$. The reason for stopping is either censoring or exit; in the latter case $z_{i,\tau_i+1} = 0$. Note that

$z_{it} = 1$ for $t \leq \tau_i$, while $z_{i,\tau_i+1} = 1$ (the firm is censored) or $z_{i,\tau_i+1} = 0$ (the firm has exited). The last observed value of z_{it} is at $t = \min(\tau_i + 1, \overline{T}_i)$.

By a standard factorization (see Billingsley, 1986), the log probability density function of $y_i = (\mathbf{y}_{i1}, ..., \mathbf{y}_{i\tau_i}, \tau_i = k, z_{i,\min(\tau_i+1,\overline{T}_i)} = j)$ can be written as:

$$\begin{aligned}
& \ln \Pr(z_{i2} = 1, ..., z_{ik} = 1, z_{i,k+1} = j | \mathbf{y}_{i1}, ..., \mathbf{y}_{ik}) + \ln f(\mathbf{y}_{i1}, ..., \mathbf{y}_{ik}) \\
= & \sum_{t=1}^{\overline{T}_i} (z_{it} \ln \Pr(z_{it} = 1 | \mathbf{y}_{i,t-1}, z_{i,t-1} = 1) + (1 - z_{it}) \ln \Pr(z_{it} = 0 | \mathbf{y}_{i,t-1}, z_{i,t-1} = 1)) \\
& + \sum_{t=1}^{\overline{T}_i} z_{it} \ln f(\mathbf{y}_{it} | \mathbf{y}_{i1}, ..., \mathbf{y}_{i,t-1}) + \ln f(\mathbf{y}_{i1}) \\
\simeq & \sum_{t=1}^{\overline{T}_i} \left(z_{it} \ln P(z_{it} = 1 | \; K_{i,t-1}, \widehat{\pi}_{i,t-1}) + (1 - z_{it}) \ln P(z_{it} = 0 | \; K_{i,t-1}, \widehat{\pi}_{i,t-1}) \right) \\
& + \sum_{t=1}^{\overline{T}_i} z_{it} \ln f(\mathbf{y}_{it} | \mathbf{y}_{i1}, ..., \mathbf{y}_{i,t-1}) + \ln f(\mathbf{y}_{i1}) \\
\equiv & \ln g(y_i; \psi)
\end{aligned} \tag{34}$$

where $f(\mathbf{y}_{i1}, ..., \mathbf{y}_{ik})$ is the density of $(\mathbf{y}_{i1}, ..., \mathbf{y}_{ik})$ corresponding to the approximate linear model (32) when k is fixed, i.e., *not* a stopping time, and the approximation in the third equation reflects that $\Pr(z_{it} = 0 | \mathbf{y}_{i,t-1}, z_{i,t-1} = 1)$ has been replaced by $P(z_{it} = 0 | \; K_{i,t-1}, \widehat{\pi}_{i,t-1})$ defined in (27).

To calculate $\ln f(y_{it} | y_{i1}, .., y_{i,t-1})$ we cast our model in a state space form with $\mathbf{y}_{it}$ as the observation vector and a_{it} as the state variable, and use the one-step ahead predictions and the predicted variances of the state variable (see Shumway and Stoffer, 2000). To obtain analytical derivatives, we use a decomposition of $\ln f(\mathbf{y}_{i1}, ..., \mathbf{y}_{ik})$ which is well-known from the EM-algorithm; see Koopman and Shephard (1992).

The quasi-likelihood estimator is given by

$$\widehat{\psi}_N = \arg\max_{\psi} L(\psi; \overrightarrow{y}_N), \tag{35}$$

where $\overrightarrow{y}_N = \{y_1, .., y_N\}$ and

$$L(\psi; \overrightarrow{y}_N) = \sum_{n=1}^{N} \ln g(y_i; \psi). \tag{36}$$

6 Estimation of structural parameters by indirect inference

Parameters and simulations from the structural model Let θ^0 denote the vector of the true parameter values of the structural (data-generating) model and ψ^* the vector of pseudo-true parameters in the quasi-likelihood (cf. (34)); i.e., the probability limit of $\widehat{\psi}_N$. From (35)-(36), ψ^* is determined by the asymptotic first-order condition

$$E_{\theta^0}\left(\frac{\partial \ln g(y_i;\psi^*)}{\partial \psi}\right) = 0,$$

where the notation $E_\theta(\cdot)$ means that the expected value refers to the data-generating (structural) model evaluated at the parameter value θ.

In indirect inference the purpose of simulations is to establish a link between θ^0 and ψ^*, which enables estimation of θ^0 from $\widehat{\psi}_N$. To this end we simulate S sequences y_i for each of the N firms, i.e., SN sequences in total. Let $y_i^{(s)}(\theta)$ denote an arbitrary simulated sequence for firm i. We will now show how $y_i^{(s)}(\theta)$ can be simulated for given θ.

We start by listing the structural parameters, θ, needed to carry out the simulations. First, the parameters needed to simulate $\mathbf{y}_{it}$ given $K_{i,t-1}$ and a_{it} are $\widetilde{\theta}_1$, θ_2 $\gamma\theta_1$, d_t and Σ_e, see (32). To simulate a_{it} we need φ and σ_a^2. To simulate π_{it} we use

$$\ln \pi_{it} = \widehat{b}_{it} + \theta_2 \ln \widehat{c}_{it} + d_t + \widetilde{\theta}_1 a_{it}. \tag{37}$$

The variables denoted by$\widehat{}$ are fixed during the estimation, with $\widehat{b}_{it}$ and $\widehat{c}_{it}$ given in (31) and (30), respectively. The simulated sequence π_{it} then follows from the simulated a_{it}. To simulate I_{it} given $K_{i,t-1}$ and π_{it} we need to evaluate $v(S_t,1)$, which requires the parameter s and also the parameters in the distribution of $g(dS_{t+1}|S_t,I_t)$. This distribution is determined by the relations

$$K_{it} = (1-\delta)K_{i,t-1} + I_{it}$$

and

$$\ln \pi_{it} - \widehat{b}_{it} - \theta_2 \ln \widehat{c}_{it} - d_t = \widetilde{\theta}_1 a_{it} = \varphi\widetilde{\theta}_1 a_{i,t-1} + \widetilde{\theta}_1 \eta_{it} = \varphi(\ln \pi_{i,t-1} - \widehat{b}_{i,t-1} - \theta_2 \ln \widehat{c}_{i,t-1} - d_{1,t-1}) + \widetilde{\theta}_1 \eta_{it}$$

where we have used (24). Under our assumption that $\ln \pi_{it}$ is an AR(1) process, which is the simplest process consistent with the assumption of a Markovian state process in

Section 3, we have

$$\ln \pi_{it} = \varphi \ln \pi_{i,t-1} + \mu + \zeta_{it}$$

where

$$\begin{aligned} \ln \pi_{i1} &= \widehat{b}_{it} + \theta_2 \ln \widehat{c}_{it} + d_t + \widetilde{\theta}_1 a_{i1} \\ \mu &= E\left(\widehat{b}_{it} - d_t - \varphi(\widehat{b}_{i,t-1} - d_{1,t-1})\right) \\ \zeta_{it} &\sim N(0, \sigma_\zeta^2). \end{aligned} \tag{38}$$

Finally, we simulate z_{it} from $\Pr(z_{it} = 0 | S_{it}, z_{i,t-1})$, which requires the parameters τ, s and ν.

While the parameters $d_1, ..., d_T, \mathbf{\Sigma}_e$ are needed to simulate $\boldsymbol{y}_{it}$, these are nuisance parameters that do not enter the structural model. Hence we will keep these parameters fixed at their quasi-likelihood estimates during the simulations and therefore they will not be considered as structural parameters. Moreover, μ and σ_ζ^2 follow trivially from (38) and φ, and hence may just be "recalibrated" by simulations of a_{it} and π_{it}; they are not considered as structural parameters that must be estimated. Therefore, the structural parameters that must be estimated by indirect inference are:

$$\theta = (\widetilde{\theta}_1, \theta_2, \gamma\theta_1, \varphi, \sigma_a^2, \tau, s, \nu).$$

The algorithm for generating an arbitrary simulated sequence $y_i^{(s)}(\theta)$ can be summarized as follows:

Let θ, $d_t = \widehat{d}_{t1}$ and $\Sigma_e = \widehat{\Sigma}_e$ be given.
If $t = 1$:

1. Let $K_{i0}^{(s)} = K_{\ i0}$ (the actual initial value of firm i)
2. draw $a_{i1}^{(s)}$ from (24)
3. draw $\pi_{i1,}^{(s)}$ from (37)
4. set $z_{i1}^{(s)} = 1$
5. Draw $\mathbf{e}_{it}^{(s)}$ from (33) and obtain $\mathbf{y}_{it}^{(s)}$ using (32).

If $t > 1$:

1. Given $K_{i,t-1}^{(s)}, \pi_{i,t-1,}^{(s)}, a_{i,t-1}^{(s)}$ and $z_{i,t-1}^{(s)} = 1$

2. Simulate $a_{it}^{(s)}$ from (24)

3. Simulate $\pi_{it}^{(s)}$ from (37)

4. Solve (18) and find $I_{it}^{(s)}$

5. Draw $z_{it}^{(s)}$ from (19)

6. If $z_{it}^{(s)} = 0$ or $t = \overline{T}_i$: stop

7. If $z_{it}^{(s)} = 1$:

   - set $K_{it}^{(s)} = (1-\delta)K_{i,t-1}^{(s)} + I_{it}^{(s)}$
   - draw $\mathbf{e}_{it}^{(s)}$ from (33) and obtain $\mathbf{y}_{it}^{(s)}$ from (32)
   - set $t = t+1$, and go to 1.

There are two challenges to the simulations: i) values for q_{Lit} for $\tau_i \leq t \leq \overline{T}_i$ are also needed (as $\tau_i^{(s)}$ may be larger that τ_i), and ii) the simulated sequence $y_i^{(s)}(\theta)$ is not continuous in θ. We can handle i) trivially by estimating a transition density $Q(q_{iLt}|q_{Li,t-1},t)$ from the data and augment the data of firm i . Regarding ii), the simulation of $\{a_{it}^{(s)}, \pi_{it}^{(s)}, \mathbf{e}_{it}^{(s)}\}_{t=1}^{\overline{T}}$ can be reduced to continuous transformations (in θ) of simulated random draws from an $\mathcal{IN}(0,1)$ distribution. Estimation of the model is done by keeping these simulated draws unchanged as θ takes on different values during the iterative estimation algorithm. This argument does not apply to $z_{it}^{(s)}$, which may change value discontinuously from 0 to 1 or vice versa as θ varies. To overcome this obstacle we replace $z_{it}^{(s)}$ by

$$E_\theta(z_{it}^{(s)}|\{a_{it}^{(s)}, \pi_{it}^{(s)}\}_{t=1}^{T}) = \Pr(z_{it}^{(1)} = 1|\{a_{it}^{(s)}, \pi_{it}^{(s)}\}_{t=1}^{T}) = \prod_{k=2}^{t} \Pr(z_{ik}^{(s)} = 1|S_{ik}^{(s)}, z_{i,k-1}^{(s)} = 1),$$

where $S_{it}^{(s)} = (\pi_{it}^{(s)}, K_{i,t-1}^{(s)})$. This can be seen as carrying out an "infinite" number of simulations of $z_{it}^{(s)}$ for each simulated sequence $\{a_{it}^{(s)}, \pi_{it}^{(s)}\}_{t=1}^{\overline{T}}$.

The efficient method of moments For any vector x and weighting matrix Ω, let $||x||_\Omega \equiv x'\Omega x$. We obtain estimates for the structural parameters by using the efficient method of moments, that is, the estimator is the solution to:

$$\widehat{\theta}^{N,S} = \arg\min_{\theta} \left\| N^{-1/2} \frac{\partial}{\partial \psi} L(\widehat{\psi}_N; \overrightarrow{y}^{(s)}_{SN}(\theta)) \right\|_{(\widehat{I}_N)^{-1}}, \tag{39}$$

where $\overrightarrow{y}^{(s)}_{SN}(\theta) = \{y_i^{(s)}(\theta)\}_{i=1}^{SN}$ is the $S \times N$ simulated sequences $y_i^{(s)}(\theta)$, with S chosen to keep the estimation uncertainty arising from simulations (i.e., the Monte Carlo standard error) below a desired tolerance level. Furthermore, $\widehat{I}_N^{-1}$ is a consistent estimator of the optimal weighting matrix I^{-1} (see Gourieroux et al., 1993). Note that some of the parameters occur both in θ and ψ, for example θ_2. Then $\widehat{\theta}_2 \neq \widehat{\theta}_2{}^{N,S}$ and, if the quasi-likelihood estimator is inconsistent, $\theta_2^* \neq \theta_2^0$.

The simulation of $y_i^{(s)}(\theta)$ was considered in the previous subsection and as shown there, the objective function in (39) is a smooth function of θ. To obtain standard errors of the indirect inference estimator, we utilize a property of the "Third Version of the Indirect Estimator" in Appendix 1 in Gourieroux et al. (1993). Here it is show that as N becomes large

$$Var(\widehat{\theta}^{N,S}) \simeq N^{-1}(1+\frac{1}{S}) \left[\frac{\partial b(\theta^0)}{\partial \theta}\right]^{-1} J^{-1} I J^{-1} \left[\frac{\partial b(\theta^0)}{\partial \theta}\right]^{-1\prime}, \tag{40}$$

where

$$b(\theta) = \arg\max_{\psi} E_\theta(\ln g(y_i;\psi)), \tag{41}$$

and

$$I = \lim_{N\to\infty} Var_{\theta^0}\left(N^{-1/2}\frac{\partial}{\partial\psi}L(\psi^*;\overrightarrow{y}_N)\right)$$
$$J = -\mathrm{p}\lim_{N\to\infty} N^{-1}\frac{\partial^2}{\partial\psi\partial\psi'}L(\psi^*;\overrightarrow{y}_N).$$

To estimate $Var(\widehat{\theta}^{N,S})$, $\partial b(\theta^0)/\partial\theta$ is obtained by finite differencing using (41), whereas I and J are obtained from the quasi-likelihood maximization (with θ^0 and ψ^* replaced by $\widehat{\theta}^{N,S}$ and $\widehat{\psi}_N$).

7 Results

7.1 Estimates of structural coefficients

In the empirical model, $\gamma\theta_1$ is the coefficient of lagged capital, $\ln K_{i,t-1}$, in the equations for $\mathbf{y}_{it}$, see (32). We can identify this (composed) coefficient, which, due to the log-linear

Table 2: **Estimates of coefficients.** Standard errors in parentheses

Industry	Directly identified coefficients				Derived estimates assuming:				
					$e = \infty$		$\varepsilon + \gamma = 1$		
	$\gamma\theta_1$	θ_2	ϱ	s	ε	γ	ε	γ	e
Wood products	.22 (.07)	−.56 (.12)	.26	.82 (.02)	.35	.10	.79	.21	1.63
Metal products	.19 (.05)	−.39 (.11)	.35	.74 (.02)	.29	.12	.74	.26	1.57
Electrical eq	.21 (.08)	−.77 (.21)	.28	.73 (.05)	.43	.08	.83	.17	2.08
Machinery	.17 (.08)	−.30 (.11)	.22	.72 (.05)	.25	.07	.81	.19	1.36
Transport eq.	.25 (.09)	−1.15 (.32)	.37	.78 (.04)	.56	.12	.79	.21	2.24
Total manuf.	.23 (.05)	−.45 (.12)	.31	.79 (.02)	.43	.16	.82	.18	1.64

form of our model, is the elasticity of a component in operating surplus (revenue, material costs or labor costs) with respect to the capital stock. The estimates of $\gamma\theta_1$ are depicted in the first column of Table 2, and they vary between 0.17 and 0.25. For pooled estimate for total manufacturing is 0.23. Hence, operating surplus increases by about 0.25 percent if the stock of capital increases by 1 percent and variable factors of production are optimized.

The low values of $\gamma\theta_1$ are in contrast to Cooper and Haltiwanger (2006), which finds an elasticity of operating surplus with respect to the stock of capital of 0.59 for a selection of US manufacturing firms. The difference may reflect that Cooper and Haltiwanger (2006) examine other types of sectors than we do; they focus on sectors with homogeneous products, for example, raw cane sugar and motor gasoline. More homogeneous products mean that the demand elasticity e in (1) is high, which, using (7), implies that $\gamma\theta_1$ is high. The intuition is straight forward: Increased production due to more capital has two counteracting effects on revenues: (i) more is produced, which, *cet. par.*, increase revenues, but (ii) the output price has to be lowered in order to sell more goods to the customers. With homogeneous products, the magnitude of (ii) is small.

The different results with respect to the curvature ($\gamma\theta_1$) may also reflect that in the Cooper and Haltiwanger study observations with negative operating surplus are removed because the authors "suspect that this [negative operating surplus] largely reflects measurement error" – this is in contrast to our approach. The estimates of the adjustment cost parameter s are in the range of 0.72 to 0.82, implying about 20 – 30 percent discount on capital in the second-hand market. Our results indicate moderate adjustment costs of capital.

As mentioned in Section 2, perfect competition can be seen as a special case of our model. We obtain perfect competition by letting the demand elasticity e in (1) approach infinity. For this limiting case we have $\theta_2 = -\varepsilon/(1-\varepsilon)$ and $\gamma\theta_1 = \gamma/(1-\varepsilon)$, see the discussion in Section 5. Hence, we now obtain an estimate of ε, and this estimate varies between 0.25 and 0.56, see Table 2. We also obtain an estimate of γ, which varies from 0.07 to 0.16. The estimate of the long-run scale elasticity $\varepsilon + \gamma$ is in the range of 0.3 to 0.7, which is much lower than most estimates of this scale elasticity – they are typically around one or somewhat higher than one. We believe our low estimate reflects that the imposed assumption of competitive market is not valid, see the discussion in Section 2.

Another special case is that of imposing a long-run scale elasticity of one. Then the estimates of the short-run scale elasticity ε lie around 0.8, see Table 2, which is close to the ratio between labor costs and value added in our data set, see Table 1. In this special case we also obtain an estimate of the demand elasticity e, which varies from 1.4 to 2.2 (1.6 for total manufacturing) – this is consistent with a high degree of market power and the high profit shares reported in Table 1.

The elasticity of substitution between labor and materials is $1-\varrho$. Because the estimates of ϱ in Table 2 lie between 0.2 and 0.4, the estimates of the elasticity of substitution are rather large. These estimates may be plausible as they are roughly in line with the corresponding parameters in the large scale computable general equilibrium model of the Norwegian economy MSG, see Bye et al. (2006).

All the coefficients in Table 2 are highly significant. Our model is parsimoniously parameterized relative to the amount of data, and we get a high goodness of fit as measured by (pseudo) R^2, which varies between 90 and 92 percent for the different industries.[10]

7.2 Exit probabilities

We now consider the estimates of the approximative exit model. These are obtained by estimating the logit equation (27). In Section 4 we derived that $\partial(v(S_t,1)-v(S_t,0))/\partial K_{t-1} \geq 0$. That is, if the capital stock increases it becomes, *cet. par.*, more valuable to continue production relative to exiting, and thus the probability to exit decreases. Hence

[10] The pseudo R^2 is defined as

$$R^2 = 1 - \frac{\text{tr } \widehat{Var}(\mathbf{e}_{it})}{\text{tr } \widehat{Var}(\widehat{\mathbf{y}}_{it} - \widehat{\mathbf{d}}_t)}$$

where tr denotes the trace, that is, the sum of the diagonal elements.

the coefficient β_K of $K_{i,t-1}^{\gamma_K}$ in (27) should be negative. In Section 4 we also derived that $\partial v(S_t, 1)/\partial \pi_t \geq 0$. Thus the coefficient β_π of $\pi_{it}^{\gamma_\pi}$ in (27) should also be negative; improved profitability lowers the probability to exit.

Table 3: **Exit probability estimates.** Standard errors of estimation in parentheses

Industry	β_K (coeff. of $K_{i,t-1}^{\gamma_K}$)	β_π (coeff. of $\pi_{it}^{\gamma_\pi}$)	γ_K	γ_π
Wood products (16)	-1.01 (.23)	-2.11 (.20)	.12 (.04)	.35 (.05)
Metal products (25)	-1.08 (.47)	-3.19 (.65)	.11 (.04)	.22 (.05)
Electrical eq. (27)	-.44 (.21)	-3.04 (.99)	.10 (.05)	.18 (.09)
Machinery (28)	-1.24 (.55)	-1.86 (.21)	.08 (.01)	.22 (.04)
Transport eq. (29-30)	-1.05 (.55)	-4.04 (2.02)	.09 (.06)	.17 (.10)
Total manufacturing	-1.17 (.33)	-2.65 (.35)	.11 (.03)	.23 (.02)

Table 4: **Elasticity of exit probabilities w.r.t. profitability and the stock of capital.**

Industry	Mean of π_{it}	Mean of* $(r+\delta)K_{it}$	Mean of Pr(exit)	Elasticity of Pr(exit) with respect to:** $K_{i,t-1}$	π_{it}
Wood products (16)	1.32	1.22	.035	-.12	-.60
Metal products (25)	2.17	1.25	.046	-.11	-.33
Electrical eq. (27)	1.85	.86	.032	-.04	-.30
Machinery (28)	1.23	.65	.035	-.10	-.34
Transport eq. (29-30)	4.24	4.10	.041	-.15	-.20
Total manufacturing	3.02	2.52	.040	-.14	-.25

*In millions EUR

** Calculated for a representative firm, i.e. with mean values of π_{it} and K_{it}

Table 3 shows that the estimate of the capital coefficient, β_K, is significant and negative in all industries as well as for the pooled industries (total manufacturing). The pooled estimate of -1.17 for β_K and 0.11 for γ_K are highly significant, with t-values above 3. Thus, the net effect of more capital is to reduce the exit probability; this is in accordance with the structural model.

The second to the last column in Table 4 shows the elasticity of the probability to exit with respect to $K_{i,t-1}$, that is, the percentage impact on the exit probability of a firm with mean values of the explanatory variables when, hypothetically, the stock of capital increases by one percent. This elasticity varies between -0.04 and -0.15. The estimated elasticity for total manufacturing is -0.14.

The third column of Table 3 shows that π_{it} (our measure of short-run profitability) has a significant negative impact on the probability to exit: the estimated value of β_π is significant and negative in all industries, and varies little across industries, ranging from -4.04 in Transport equipment to -1.86 in Machinery, with a common estimate of -2.65 for total manufacturing. The exponent of the corresponding power function, γ_π, is estimated to be around 0.2 in most industries. The lowest estimate is found in Transport equipment (0.17) and the highest in Wood products (0.35). When all industries are pooled, the estimates of β_π and γ_π are -2.65 and 0.23, respectively. Both estimates are significantly different from zero; the t-value corresponding to β_π is 8, whereas the t-value corresponding to γ_π is 20.

The last column in Table 4 shows the elasticity of the exit probability with respect to π_{it} (for given capital stock, $K_{i,t-1}$), that is, by how many percent the exit probability changes – for a firm with mean values of the explanatory variables – when, hypothetically, our measure of short-run profitability (π_{it}) increases by one percent. The Table shows that this elasticity varies across sectors from -0.20 in Transport equipment to -0.60 in Wood products, with a common estimate for total manufacturing of -0.25. This indicates that the impact on exit of improved profitability (measured by π_{it}) is stronger than the impact of more capital, see discussion above.

In order to evaluate the aggregate performance of our model, in each year we divide firms into two groups: *closing-down firms* in year t are those that exited during $t+1$, and *non-exiting* firms are those that did not exit during the whole observation period. Our two definitions imply that firms exiting in $t+1+p$ ($p \neq 0$) are not included in any of the two groups in year t. For each firm we are able to estimate – for each (relevant) year – the probability to exit in the next year.

Figure 4 plots annual averages of the estimated exit probability of the two groups of firms. Our model discriminates between the two categories of firms: the exit probabilities of the closing-down firms are generally higher than those of the non-exiting firms. However, the differences between the two groups vary a lot over time (for a given industry), with large differences in some years and quite small ones in others.

Because our estimator maximizes the fit of the model to the observed data, the result that closing-down firms have a higher estimated exit probability than non-exiting firms

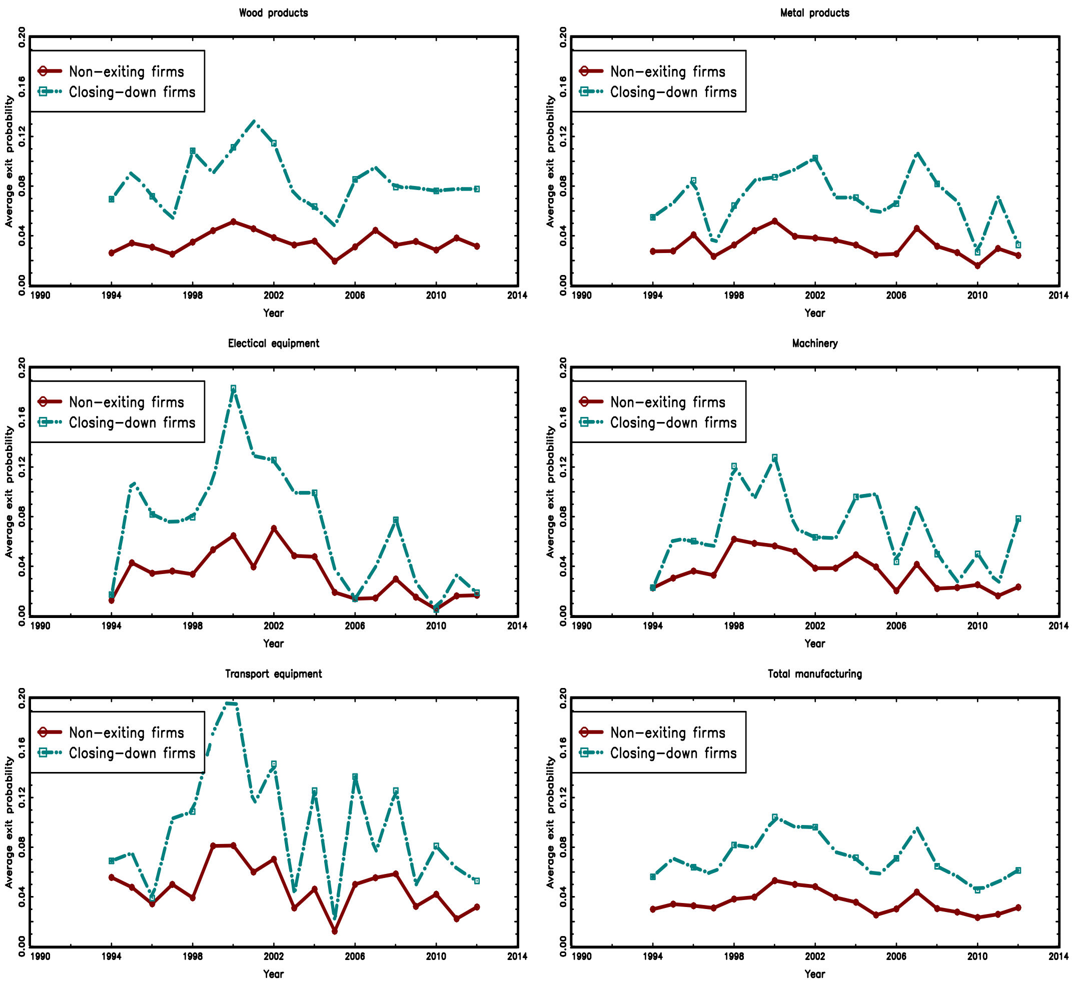

Figure 4: Estimated aggregate exitprobabilities for non-exiting and closing-down firms in $t = 1994, ..., 2012$

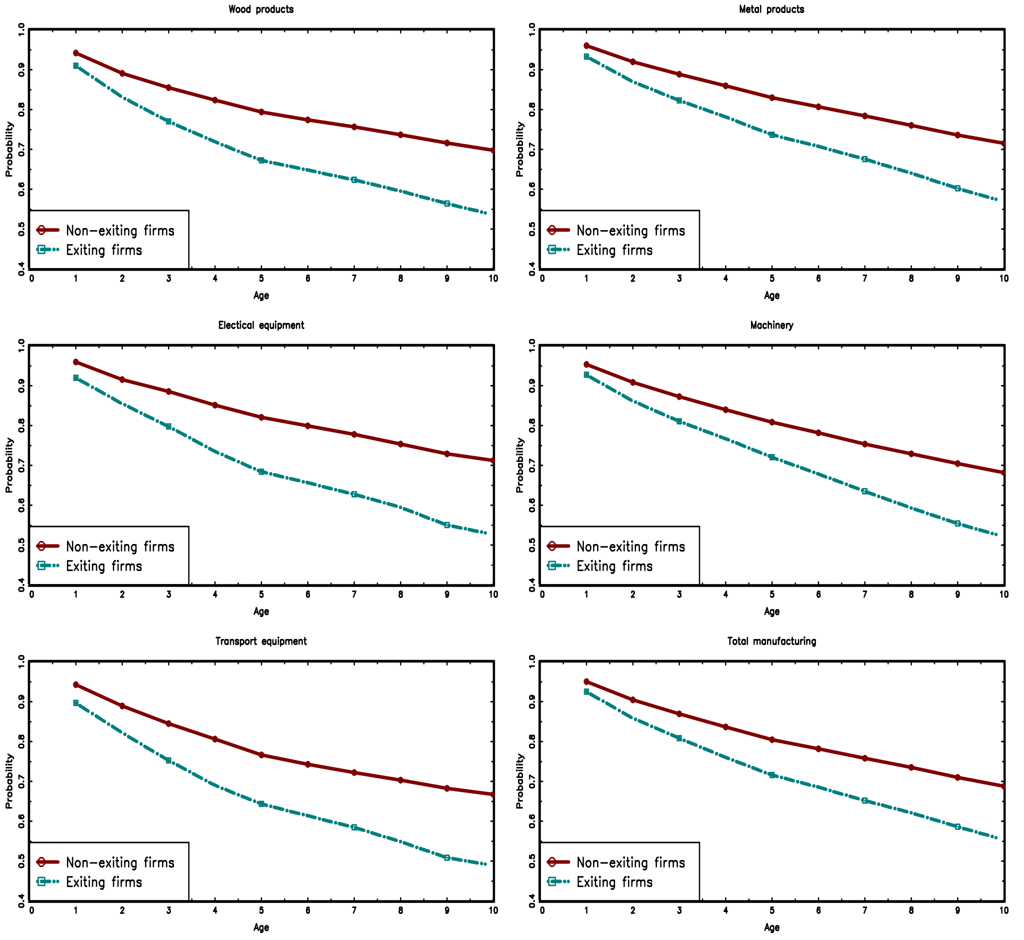

Figure 5: Estimated survival probabilities as a function of firm age for non-exiting and exiting firms

may seem uninteresting. To understand our result, consider the hypothetical case in which all the realized variables of the firms, including exit, are assigned in a purely random way. Then, by assumption, there is no relation between exit and the covariates, that is, short-run profitability and the size of the capital stock. Hence, in this constructed data set profitability and the stock of capital have no impact on the estimated probability to exit, and the estimated coefficients will be (approximately) zero. We find that both covariates have a *significant* impact on the estimated exit probability. The estimated model implies that there is a substantial difference between the exit probability of closing-down and non-exiting firms.

7.3 Survival functions

Figure 4 illustrates the difficulty of predicting the *exit time*; the estimated exit probabilities of the closing-down firms are erratic and vary a lot over time. The interpretation of Figure 4 is, however, not straightforward as the graphs incorporate different effects. First, they reflect temporal variations in both firm-specific conditions (e.g., technological changes) and in industry-specific conditions (changes in factor prices and demand). Second, the graphs of the closing-down firms reflect a composition effect as different firms are operating in different years; if entrants to the industry (on average) have higher exit probability than the incumbents, then (*cet. par.*) the average exit probability tends to decrease over time as the share of incumbent firms will increase compared to 1995 when all firms in our sample were start-up firms.

To control for the second effect, that is, self-selection, we use our estimated model to simulate survival functions. These show the probability that a firm has survived until the end of year t as a function of time after entry and initial conditions. We construct the survival functions as follows: We impose that all firms enter in the same year, henceforth referred to as year 0. For each firm we use the estimated logit model of exit and the values of the observed variables, $(K_{i,t-1}, \widehat{\pi}_{i,t-1})$, in the first year the firm is (actually) operative to calculate the exit probability in year 0, that is, the probability to exit during the next year (year 1). A proportion of the firms is then removed by the following procedure: For each firm a random number is drawn from the uniform distribution on $[0, 1]$. Firms with a number lower than its estimated exit probability are removed. For each of the "surviving" firms, a new exit probability is estimated using the estimated model and the values of $(K_{i,t-1}, \widehat{\pi}_{i,t-1})$ in the second year the firm is contained in the data set. Then a proportion of the firms is removed, and so on. If a firm "survives" more years than in the real data set, we use the econometric model to simulate the values of $(K_{i,t-1}, \widehat{\pi}_{i,t-1})$ from our structural econometric model, see Section 6.

This experiment was repeated 100 times. In general, a firm will experience many different exit years. The frequency of exit years was used to construct the survival function of a firm as follows: Let $Z_{it}^{(s)}$ be an indicator function which is one if – in the s'th simulation – firm i has not exited by year t. Note that $Z_{it}^{(s)} = 1$ is conditional on $Z_{i0}^{(s)} = 1$ since all firms are operative at the end of year 0. By repeated simulations, $s = 1, ..., 100$, a firm-specific *conditional* survival function, $S_i(t) = P(Z_{it} = 1 | Z_{i0} = 1, \mathbf{y}_{i0})$ was estimated as

$$\widehat{S_i(t)} = \frac{1}{100} \sum_{s=1}^{100} Z_{it}^{(s)}.$$

We group firms according to whether they actually exited in the data period - these are termed exiting firms - or never exited - these are referred to as non-exiting firms. We construct survival functions for each group by averaging the survival functions over all firms in each group; see Figure 5. By comparing the survival functions for exiting and non-exiting firms, we can evaluate to what degree our model is able, *ex ante*, to "pick" the firms that actually exited during the observation period. Overall we find that our model clearly discriminates between the two groups. For example, for the total manufacturing sector we find that the ex ante survival probability of exiting firms is about 55 percent after 10 years, compared to 70 percent for non-exiting firms. Similar differences are found for the individual industries. Our results suggest that the main characteristic of an exiting firm is not that its annual exit probability is very much higher than that of a surviving firm, but rather that the difference in annual exit probabilities is highly persistent. Hence, it is the cumulated effect of higher annual exit probabilities over many years – compared with the average firm – that causes a firm to exit.

7.4 Computational aspects

Figure 6 shows the profiles of the objective function (39) for total manufacturing for two parameters: The exponent of (lagged) capital in the operating surplus function $(\gamma\theta_1)$ and the adjustment cost parameter (s). These profiles are obtained by varying one parameter at the time in a neighborhood of the final estimate, keeping all the other parameters fixed at the optimum. The objective function is close to zero at the minimum, and it appears to be smooth and (locally) almost quadratic. A similar picture emerges for the other structural parameters (not shown).

To minimize the objective function in (39) we use a robust derivative-free method: We apply a version of the conjugate direction method developed by Brent (1973, Ch. 7) in combination with the derivative-free line search algorithm from the Numerical Recipes library (Press et al., 1994, p. 419). The optimizations illustrated in Figure 6 converged after 24 hours using GAUSS on a Linux server. In contrast, it took 15 minutes to estimate the auxiliary model.

The results reported in the tables use the corresponding quasi-likelihood estimates as starting values for the parameters that enter both the data generating and auxiliary model. Generally, we found that the indirect inference estimates are robust to the choice of starting values, although extreme starting values can make the algorithm converge towards the boundary of the parameter space.

8 Conclusion

In the Introduction we raised three questions: Is there a relationship between firm exit and profitability, what causes firms to exit, and what are the characteristics that distinguish exiting firms from non-exiting firms? Using a structural econometric model we have derived explanatory variables from economic theory and estimated models for five Norwegian manufacturing industries separately. The results show that when exit is defined as a state in which production at the site has come to a permanent stop, increased profitability significantly lowers the exit probability, or put differently; low profitability causes firms to exit. We have also found a clear difference in the estimated exit probabilities between firms that exited in the sample period (1994-2012) and firms that did not exit. According to our study, exiting firms differ from non-exiting firms as their annual exit probabilities are *persistently* higher. On the other hand, exiting firms are not characterized by having a very high exit probability just prior to exit, which reflects that there are no (negative) profitability shocks in the last years prior to exit.

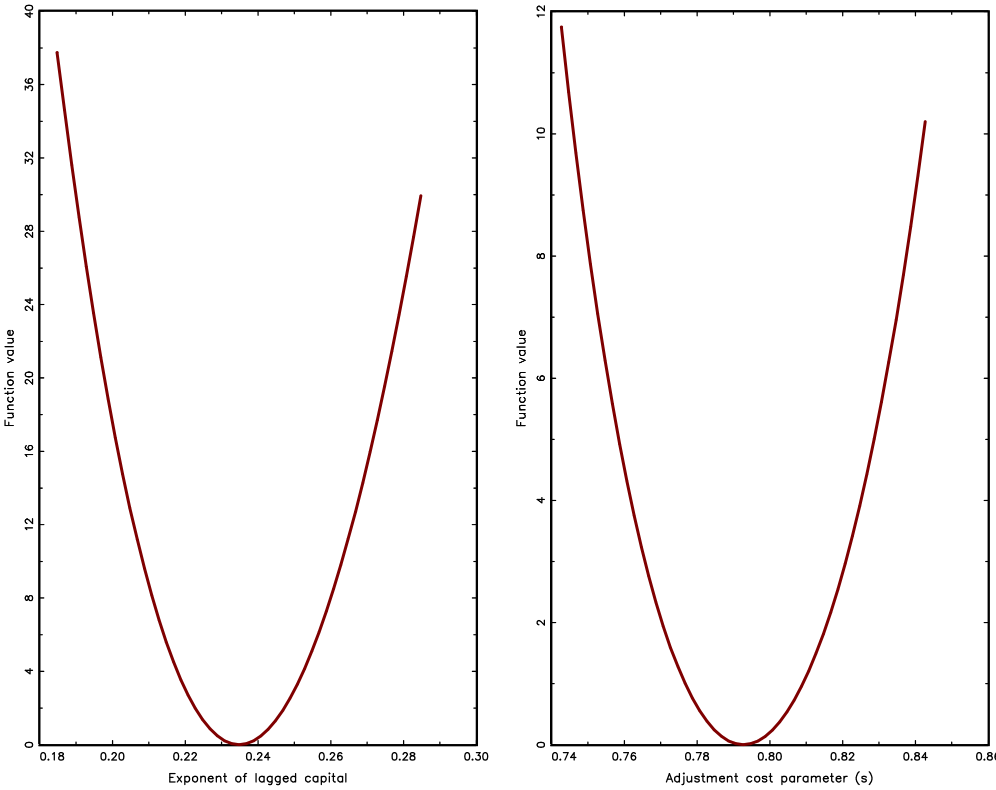

Figure 6: Profiles of the indirect inference objective function for the parameters s and $\gamma\theta_1$.

In Section 2 we argued that our data suggest a weak relationship between profitability and exit. Yet, our estimation results clearly indicate such a relationship. We believe this shows the power of econometric modeling and methods. For example, in our model cost of adjustment, which reflects that the resale price of capital is lower than the purchasing price of new capital, is potentially a key factor in determining exit. But this type of cost is not included in our capital data. However, by deriving an expression for the adjustment cost of capital we have been able to incorporate this factor into the expression for the probability of firm exit and have demonstrated both theoretically and empirically that the probability to exit is a decreasing function of the capital stock.

In the Introduction we referred to some articles that have documented empirically that the age of a firm is related to its probability to exit. In our derived logit model, age is not an explanatory variable simply because age plays no role in the theoretical foundation of the econometric model. On the contrary, age is a function of survival, and hence a highly endogenous variable – not a control variable. However, age (survival) may play a role through learning processes, see Jovanovic (1982) and Pakes and Ericson (1998), and hence have a separate (causal) effect on the exit probability. Hence, one extension of our model would be to incorporate learning effects.

References

[1] Abel, A and J. Eberly (1994): A Unified Model of Investment under Uncertainty. *American Economic Review*, 84,1369–84.

[2] Abel, A. and J. Eberly (1996): Optimal investment with costly reversibility. *Review of Economic Studies*, 63, 581–593.

[3] Acemoglu,D, U.Akcigit, N. Bloom, W .R. Kerr (2013): Innovation, Reallocation and Growth. NBER Working Paper No. 18993.

[4] Agarwal, R. and D. B. Audretsch (2001): Does entry size matter? The impact of the life cycle and technology on firm survival. *Journal of Industrial Economics*, XLIX (1), 21-43.

[5] Andersen, T.G., H.J. Chung and B.E. Sørensen (1999): Efficient method of mo-

ments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics,* 91, 61–87.

[6] Andersen, T.G. and J. Lund (1997): Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics,* 77, 343–378.

[7] Asphjell, M.K., W. Letterie Ø.A. Nilsen and G.A. Pfann (2014): Sequentiality Versus Simultaneity: Interrelated Factor Demand. *Review of Economics and Statistics.*96, 986-998.

[8] Bansal, R., A.R. Gallent and G. Tauchen (2007): Rational Pessimism, Rational Exuberance, and Asset Pricing Models. *Review of Economic Studies*, 74,1005-1033.

[9] Bartelsman, E., J. Haltiwanger and S. Scarpetta (2013): Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1), 305-34.

[10] Billingsley, P. (1986): *Probability and Measure.* John Wiley, New York.

[11] Bloom, N., S. Bond and J. van Reenen (2007): Uncertainty and Investment Dynamics. *Review of Economic Studies*, 74, 391-415.

[12] Brent, R.P. (1973): *Algorithms for minimization without derivatives.* Prentice-Hall, Englewood Cliffs.

[13] Bye, B., T.R. Heggedal, T. Fæhn and B. Strøm (2006): A CGE model of induced technological change. A detailed model description. Documents 2006/11, Statistics Norway.

[14] Byrnes, J.P, D.C. Miller and W.D. Schafer (1999): Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3), 367-383.

[15] Caggese, L.A. (2007). Testing financing constraints on firm investment using variable capital. *Journal of Financial Economics*, 86(3), 683–723.

[16] Caves, R. E. (1998): Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, 36, 1947–1982.

[17] Cooper, R. W. and J. C. Haltiwanger (2006): On the nature of capital adjustment costs. *Review of Economic Studies*, 73, 611–633.

[18] Croson, R. and U. Gneezy (2009): Gender Differences in Preferences. *Journal of Economic Literature*, 47 (2), 448-474.

[19] Das, S. (1992): A micro econometric model of capital utilization and retirement: The case of the cement industry. *Review of Economic Studies*, 59, 277-297.

[20] Disney, R., J. Haskel and Y. Heden (2003): Entry, exit and establishment survival in UK manufacturing. *Journal of Industrial Economics*, LI (1), 91-112.

[21] Dixit, A. K and R.S Pindyck (1994): *Investment under uncertainty.* University Press, Princeton, NJ.

[22] Doms, M., T. Dunne and M. J. Roberts (1995): The role of technology use in the survival and growth of manufacturing plants. *International Journal of Industrial Organization*, 13 (4), 523-542.

[23] Dunne, T., M. J. Roberts and L. Samuelson (1988): Pattern of firm entry and exit in U.S. manufacturing industries. *RAND Journal of Economics*, 19(4), 495-515.

[24] Foster, L., J. Haltiwanger and C. Syverson (2008): Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98 (1), 394-425.

[26] Gallant A.R. and J.R. Long (1997). Estimating stochastic differential equations efficiently by minimum chi-square. *Biometrika*, 84, 125-141.

[26] Gallant A.R. and G. Tauchen (1996). Which moments to match? *Econometric Theory,* 12, 657-681.

[27] Gourieroux C., A. Monfort and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics*, 8, Supplement, S85-S118.

[28] Hennessy, C.A. and T.M. Whited (2007): How Costly Is External Financing? Evidence from a Structural Estimation. *The Journal of Finance,* 62, 1705–1745.

[29] Hopenhayn, H (1992): Exit, selection, and the value of firms. *Journal of Economic Dynamics and Control*, 16, 621-653.

[30] Hsieh, C. T. and P. J. Klenow (2009): Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, CXXIV (4), 1403-48.

[31] Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica,* 50 (3), 649-670.

[32] Klepper, S. (2002): Firm survival and the evolution of oligopoly. *RAND Journal of Economics*, 33 (1), 37-61.

[33] Koopman, S. J and N. Shepard (1992): Exact score for time series models in state space form. *Biometrika*, 79, 823–826.

[34] Levinsohn, J. and A. Petrin (2003): Estimating production functions using inputs to control for unobservables. *Review of Economic Studies,* 70, 317–342.

[35] Marshall, A. (1966): *Principles of Economics.* 8th edition. Macmillan, London.

[36] Mata, J., P. Portugal and P. Guimarães (1995): The survival of new plants: Start-up conditions and post-entry evolution. *International Journal of Industrial Organization*, 13 (4), 459-481.

[37] Nilsen, Ø.A. and F. Schiantarelli (2003): Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities. *Review of Economics and Statistics*, 85, 1021–1037.

[38] OECD (2001): *Measurement of capital stocks, consumption of fixed capital and capital services.* OECD manual 131 (Paris and Washington).

[39] Olley, G. S. and A. Pakes (1996): The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64, 1263–1297.

[40] Pakes, A. and R. Ericson (1998): Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*, 79, 1-45.

[41] Parker, S. C. (2004): *The economics of self-employment and entrepreneurship.* Cambridge University Press.

[42] Pérez, S. E., A. S. Llopis and J. A. S. Llopis (2004): The determinants of survival of Spanish manufacturing firms. *Review of Industrial Organization*, 25, 251-273.

[43] Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1994): *Numerical recipes in C*. Cambridge University Press, Cambridge.

[45] Raknerud, A., D. Rønningen and T. Skjerpen (2004): Documentation of the Capital database. A database with data for tangible fixed assets and other economic data at the firm level. Documents 2004/16, Statistics Norway. Downloadable from: http://www.ssb.no/english/subjects/10/90/doc_200416_en/doc_200416_en.pdf

[45] Raknerud, A., D. Rønningen and T. Skjerpen (2007): A method for improved capital measurement by combining accounts and firm investment data. *Review of Income and Wealth*, 53, 397-421.

[46] Raknerud, A and Ø. Skare (2012): Indirect inference methods for stochastic volatility models based on non-Gaussian Ornstein-Uhlenbeck processes. *Computational Statistics and Data Analysis,* 56, 3260-3275.

[47] Rust, J. (1994): Structural estimation of Markov decision processes. In R.F. Engle and D.L. McFadden *(eds.) Handbook of Econometrics, Volume IV*. Amsterdam, Elsevier.

[48] Shumway, R.H. and D.S. Stoffer (2000): *Time Series Analysis and Its Applications.* New York, Springer.

[49] Stock, J.H. and D.A. Wise (1990): Pensions, the option value of work, and retirement. *Econometrica*, 58, 1151-1180.

[50] Sutton, J. (1997): Gibrat's legacy. *Journal of Economic Literature*, 35, 40–59.

9 Appendix: Proof of Proposition 1

Let $v(S_t, z_t)$ denote net present value given S_t and z_t:

$$v(S_t, z_t) \quad = \quad \max_{I_t} \left\{ u(S_t, I_t, z_t) - \Pi_t + \frac{1}{1+r} E_t \left[V(S_{t+1}, \varepsilon_{t+1}) \right] \right\}. \tag{42}$$

Then (16) follows by definition. If $z_t = 0$, t is the terminal period and $v(S_t, 0) = u(S_t, -(1-\delta)K_{t-1}, 1) - \Pi_t = -c(-(1-\delta)K_{t-1})$, which proves (17).

To prove (18), assume a finite horizon problem and let $v_T(S_t, 1)$ be defined in the same way as $v(S_t, 1)$ in (42), except that ∞ is replaced by T in the summation limit. Hence, $v_T(S_t, z_t) + \Pi_t + \varepsilon(z_t)$ is the net present value in period t of choosing z_t and then make optimal decisions with regard to I_t and $(I_{t+k},\ z_{t+k})$ for $k = 1, ..., T$, where $t + T$ is the terminal period. Thus we consider a T–period ahead problem. For example, $v_0(S_t, 1)$ is the solution of the static problem ($T = 0$), $v_1(S_t, 1)$ is the one-period ahead problem ($T = 1$), etc. Let $V_T(S_t, \varepsilon_t)$ denote the value function (16) corresponding to the T-period ahead problem: $V_T(S_t, \varepsilon_t) = \max_{z_t}[\Pi_t + v_T(S_t, z_t) + \varepsilon(z_t)]$. Obviously, $v_0(S_t, 1) = \max_{I_t}[-c(I_t)]$. Furthermore,

$$\begin{aligned} v_1(S_t, 1) \quad &= \quad \max_{I_t} \left\{ -c(I_t) + \frac{1}{1+r} \int V_0(S_{t+1}, \varepsilon) h(\varepsilon) d\varepsilon \, g(dS_{t+1}|S_t, I_t) \right\} \\ &= \quad \max_{I_t} \left\{ -c(I_t) + \frac{1}{1+r} \int \max \left\{ \Pi_{t+1} + v_0(S_{t+1}, 0) + \varepsilon(0), \Pi_{t+1} + v_0(S_{t+1}, 1) + \varepsilon(1) \right\} h(\varepsilon) d\varepsilon \times \right. \\ &\quad\quad g(dS_{t+1}|S_t, I_t)\} \\ &= \quad \max_{I_t} \left\{ -c(I_t) + \frac{1}{1+r} \int \left[\Pi_{t+1} + \frac{1}{\tau} \ln \left[\exp(-\tau c(-(1-\delta)K_t) + \gamma_0) + \exp(\tau v_0(S_{t+1}, 1) + \gamma_1) \right.\right.\right. \\ &\quad\quad g(dS_{t+1}|S_t, I_t)\}, \end{aligned}$$

where the integrand after the last equality is the so-called "social surplus" function. The last equality follows from (13) and a well-known property of the extreme value distribution (see Rust, 1994). By backward recursion we obtain:

$$\begin{aligned} v_T(S_t, 1) \quad &= \quad \max_{I_t} \left\{ -c(I_t) + \frac{1}{(1+r)} \times \right. \\ &\quad\quad \int \left[\Pi_{t+1} + \frac{1}{\tau} \ln \left[\exp(-\tau c(-(1-\delta)K_t) + \gamma_0) + \exp(\tau v_{T-1}(S_{t+1}, 1) + \gamma_1) \right] \right] \times \\ &\quad\quad g(dS_{t+1}|S_t, I_t)\}. \end{aligned} \tag{43}$$

Under the regularity conditions of Rust (1994), equation (43) defines contraction mapping so that

$$\sup_S |v_T(S, 1) - v_{T-1}(S, 1)| \to 0 \text{ as } T \to \infty.$$

Then there exists a limiting function $v(S,1)$ that satisfies (18). Finally, from (16):

$$\begin{aligned} P(z_t &= 0|S_t) = P\left(v(S_t,0)+\varepsilon(0) > v(S_t,1)+\varepsilon(1)|S_t\right) \\ &= P\left(\tau v(S_t,0)+\gamma_0+(\tau\varepsilon(0)-\gamma_0) > \tau v(S_t,1)+\gamma_1+(\tau\varepsilon(1)-\gamma_1)|S_t\right) \\ &= \frac{1}{1+\exp\{-[\tau v(S_t,0)-\tau v(S_t,1)+\gamma_0-\gamma_1]\}}, \end{aligned}$$

where we in the last equation used that $\tau\varepsilon(z)-\gamma_z$ has a standard extreme value distribution and is independent for $z=0,1$. Hence, (19) follows from (17).

Q.E.D.