Financing Constraints and State Dependent Adjustment Dynamics: A New Estimation Routine Using Micro-Level Survey Data

Ulf von Kalckreuth
Deutsche Bundesbank

To be presented at the
Conference in Honor of George von Fürstenberg
Indiana University, Bloomington, 28-29 April 2006

This version: 24.04.06

Abstract
In recent years, the traditional methodology of detecting financing constraints, pioneered by Fazzari, Hubbard and Petersen (1988), has been widely criticised. This paper departs from the received practice in several ways: We use direct evidence from survey data instead of a-priori groupings of firms. Instead of relying on differential cash flow for identification, we compare the adjustment dynamics of constrained and unconstrained firms. Most importantly, instead of grouping firms once and for all, we distinguish – for any given firm – between episodes of financially unconstrained growth, financially constrained growth and contraction. The main hypothesis is that the speed of adjustment is slower for constrained firms.

To this end, we develop a GMM estimation routine for a dynamic fixed effects panel data model data that allows for state dependent persistence. The estimator is derived and simulated. First results using real data indicate that the adjustment dynamics in constrained and non-constrained episodes can be successfully distinguished. The new estimator has a wide area of important applications, ranging from questions of optimal capital structure and investment behaviour to the price setting of firms and household finances.

Key Words: Financing constraints, investment, dynamic panel data models

JEL Classification: D210, D920, C230

Correspondence: Ulf von Kalckreuth, Deutsche Bundesbank, Economic Research Centre, Wilhelm Epstein-Str. 14, D-60431 Frankfurt am Main, Germany. E-mail: ulf.von-kalckreuth@bundesbank.de . Website: http://www.von-kalckreuth.de

The views expressed in this paper do not necessarily reflect those of the Deutsche Bundesbank. All errors, omissions and conclusions remain the sole responsibility of the author.
1. Introduction

This paper develops and applies an estimation technique which is tailor-made for testing a hypothesis on financing constraints and their effect on the accumulation dynamics. In this respect it is a companion paper to von Kalckreuth (2004), where this hypothesis is developed and first steps at testing its implications are made. But the methodology described here should have much broader applications.

Empirical work on financial constraints has traditionally been based on an approach pioneered by Fazzari, Hubbard and Peterson (1988). If the investment of supposedly financially constrained firms shows a higher sensitivity to internal finance than the investment of their supposedly unconstrained counterparts, this is seen as evidence for the existence of binding financial constraints. In the meantime, this approach has been forcefully criticised. Kaplan and Zingales (1997) state that there is no theoretical reason why – in a comparison between firms – a larger cost differential between internal and external finance might lead to a higher cash-flow sensitivity, as opposed to just comparing the extreme cases of a constrained firm and an absolutely unconstrained one. A non-monotonic relationship between the cost differential and excess sensitivity is perfectly conceivable.\(^1\) On the other hand, it has been shown theoretically that, under certain conditions, cash flow terms can be significant even in the absence of financial constraints.\(^2\) Ultimately, there is a pervasive missing variable problem. Cash flow is a close relative to profit, a summary measure of all that is important for a firm, and it is useful in predicting future values of variables relevant to the current investment decision.\(^3\)

In von Kalckreuth (2004), von Kalckreuth (2006) and this paper, we use a direct approach by relying on explicit statements by the firms themselves. We are able to explore the micro data base of the Ifo Institute’s Investment Test (Investitionstest, IT) for the manufacturing sector in western Germany during the years 1988 to 1998. During these eleven years, the autumn wave yields 25,643 observations on a total of 4,443 firms, with 2,331 firms per year on average. Apart from its size and coverage, the data set has three important characteristics that are relevant to our problem. First, it contains many small firms, on which very little information is available from micro data sets based on quoted companies. Although large firms are clearly

\(^1\) The discussion was continued in Fazzari, Hubbard and Peterson (2000) and Kaplan and Zingales (2000).
\(^2\) See the models by Abel and Eberly (2003), Cooper and Ejarque (2001), and Gomes (2001).
\(^3\) This argument is developed formally in Appendix B of Chirinko and von Kalckreuth (2002).
over-sampled, almost 50% of the IT observations refer to firms with fewer than 200 employees, and 19.5% of the firms have fewer than 50 employees. Second, firms report on their innovation behaviour. They state whether or not a product innovation was achieved during the past year and whether or not it was fundamental in a technological sense. Third, the data set contains information on financial constraints that firms face in their investment decisions. Notably, a number of firms (around 26.2% of respondents) explicitly state that their investment demand is limited by the cost and/or the unavailability of finance. Although part of this may be due to the workings of the classical interest rate channel, these aggregate effects can be eliminated by the use of time dummies, focussing on differential changes in time.

In the companion paper we have shown that a specific pattern with respect to the distribution of investment over time should be expected to hold for financially constrained firms. Given a shock, an unconstrained investor can adapt rapidly, or even instantaneously if adaptation costs are unimportant. The bulk of investment spending will take place in the first few periods, and there may be a spike in the first period. If the investor is financially constrained, however, marginal costs of finance will increase with the amount of spending, possibly to infinity. In such a setting, the investor has to equalise marginal costs of finance and the marginal value of new investment in each period. After an initial debt-financed increase in the capital stock that leads to a worsening of the financial position, the firm needs internal finance to continue the expansion and to repair balance sheets gradually. Thus, the adaptation process will be spread over time. In short: money buys time! This crucial difference in the adaptation dynamics can be used to identify financially constrained firms, or better, whether a subset of supposedly constrained firms really is, without having to take recourse to cash flow sensitivities.

It is possible to condense the dynamics of the model into a single diagram. In Figure 1, the decreasing schedule represents marginal return on investment, whereas the increasing schedule with the flat portion depicts the costs of finance. Given a profitability shock, indicated by arrow (1), the financially constrained firm will immediately invest up to the point where the costs of external finance are equal to marginal profitability of investment. The difference between marginal profitability and the costs of finance in the steady state winds up a “clockwork”. The firm retains profits, indicated by arrow (2), expanding equity base and physical capital at the same time. The “clock” winds down, along the falling schedule that depicts marginal returns on investment, until dynamic equilibrium is reached again.
The hypothesis is examined empirically by von Kalckreuth (2006), using an entirely qualitative set of survey data on UK firms. The duration of capacity restrictions was compared between firms that characterised themselves as being financially constrained, and others that did not. The Ifo Investment Test has a different and more specific informational content, in that many key variables are continuously scaled, so that the adjustment dynamics can be much more closely observed.

Von Kalckreuth (2004) proceeds to sort firms into two groups, according to whether they are predominantly financially constrained or not. Apart from endogeneity problems, this approach has a serious drawback: it does not make any use of the time variation in the financing constraints variable for a given firm. This is the most valuable sort of information in micro-econometrics, when many important aspect of the process in question are unobserved, but can be trusted to be relatively stable in time. In this case it is the variation in the left hand variable following a variation of the explanatory variable that helps to identify structural relationships.

What we are looking for are not simply variations in the left hand variable, but variations in their speed of adjustment. This is similar to the so-called threshold models in time series analysis, see, for example, Enders and Granger (1998) or Enders and Siklos (2001). There are two important differences. First, it is realistic to model the target value as an unobserved, firm specific variable. On the other hand, we have indicator variables that inform us on the regime in which the adjustment process is situated, so that no threshold needs to be estimated.

It is important to note that the regime variable is endogenous: Financing constraints will partly depend on the financing needs. We might think of modelling the interaction of financing constraints and capital accumulation as an interdependent system, using a ML method, see...
e.g. Maddala (1983) Ch. 8. This is not feasible here, as the most important explanatory variable for financing constraints is missing, namely the financial structure. Therefore, we want to isolate the line of causality that runs from financing constraints to the speed of adjustment by using a GMM approach. To the best of our knowledge, regime dependent dynamic adjustment has not been studied in a GMM context before.

This paper develops two sets of moment conditions for the use of GMM estimators, by adapting the principles underlying the widely used estimators for linear dynamic panel data models for the use in our non-linear problem. We believe that this estimation technique – being very much related to the everyday-tools of dynamic micro-econometrics and not difficult to implement – can be fruitfully applied to a large number of other problems.

2. A regime specific adjustment process

We want to model a process that reverts to a static equilibrium value characteristic for individual \( i \), and not necessarily identical for all observational units. For individual \( i \), a number of \( n+1 \) different regimes is defined, \( r_{i,t} \in \{0,1,2,...,n\} \). The speed of adjustment differs according to the regime:

\[
\Delta y_{i,t} = -\alpha_{i,t-1}(y_{i,t-1} - \mu_i) + \varepsilon_{i,t},
\]

with an adjustment coefficient \( \alpha_{i,t-1} \) that depends on the regime \( r_{i,t-1} \).

\[
\alpha_{i,t-1} = \alpha(r_{i,t-1}) = \begin{cases} 
\alpha_0 & \text{for } r_{i,t-1} = 0 \\
\alpha_1 & \text{for } r_{i,t-1} = 1 \\
\alpha_n & \text{for } r_{i,t-1} = n 
\end{cases}
\]

The regime is assumed to be observed, e.g. by means of survey responses. It may be defined by a firm being financially constrained or not, or expanding or not. Without any loss of generality, we will concentrate on the case of two regimes: \( r_{i,t} \in \{0,1\} \). Adding \( y_{i,t-1} \) to both sides gives

\[
y_{i,t} = (1-\alpha_{i,t-1})y_{i,t-1} + \alpha_{i,t-1}\mu_i + \varepsilon_{i,t}
\]

The second component is not a fixed effect, as it would be in the model of regime independent adjustment, but it is an individual specific stochastic process. With respect to the error process, we will start by assuming that it is serially uncorrelated. Furthermore, it is assumed to be correlated with the current regime indicator, \( r_{i,t} \). The regime indicator follows a process that has two important characteristics: it has a finite memory, and it is correlated with current, but not with past realisations of \( \varepsilon_{i,t} \). We will assume it to be driven by a moving average process of order \( q \):
\[ z_{t,j} = \sum_{j=0}^{q} b_j \eta_{t-j}, \quad E\eta_{t,j}^2 = \sigma^2_\eta, \quad E\eta_{t,j}\eta_{t-k} = 0 \forall k > 0, \quad \text{with} \] (3)

\[ E\eta_{t,j}\epsilon_{t,j} = c, \quad E\eta_{t,j}\epsilon_{t-k} = 0 \forall k > 0 \text{ and } r_{t,j} = \text{Ind}(z_{t,j} \geq 0). \]

For later use, we want to derive the backward solution \( y_{t,j} \), as a function of current and past values of \( \epsilon_{t,j}, \ r_{t,j} \) (or equivalently \( \alpha_{t,j} \)) and \( \mu_i \). In \( t = 0 \), we assume to be given \( z_{t,0} \) (by implication also \( r_{t,0} \) and \( \alpha_{t,0} \), \( \eta_{t,0} \) and \( y_{t,0} \)). By repeated substitution, we find the following

**Proposition 1**: For \( t \geq 2 \), the backward solution to the adjustment equation is given by:

\[ y_{t,j} = \left[ y_{t,0} - \mu_i \right] \prod_{k=0}^{t-1} (1 - \alpha_{t,k}) + \sum_{l=1}^{t-1} \prod_{k=l}^{t-1} (1 - \alpha_{t,k}) \epsilon_{t,j} + \mu_i + \epsilon_{t,j} \] (4)

The first term gives the effect of the initial conditions. The second term gives the accumulated effects of the innovations \( \epsilon_{t,j} \), and the third term is the individual specific shifter. As \( \alpha_{t,j} \) and \( \epsilon_{t,j} \) are contemporaneously correlated, the expected value of the second term is not zero.

**Proof**: 1) The proposition holds for \( t = 2 \). 2) If it holds for \( t-1 \), it also holds for \( t \\
\[ y_{t,j} = \left[ y_{t,0} - \mu_i \right] \prod_{k=0}^{t-1} (1 - \alpha_{t,k}) + \sum_{l=1}^{t-1} \prod_{k=l}^{t-1} (1 - \alpha_{t,k}) \epsilon_{t,j} + \mu_i + \epsilon_{t,j} \]

\[ = \left[ y_{t,0} - \mu_i \right] \prod_{k=0}^{t-1} (1 - \alpha_{t,k}) + \sum_{l=1}^{t-1} \prod_{k=l}^{t-1} (1 - \alpha_{t,k}) \epsilon_{t,j} + \mu_i + \epsilon_{t,j} \]

For reference purposes, it is useful to expose the structure of the term that captures the effect of innovations:

\[ \sum_{l=1}^{t-1} \prod_{k=l}^{t-1} (1 - \alpha_{t,k}) \epsilon_{t,j} = \epsilon_{t,j} \left(1 - \alpha_{t,1}\right) \left(1 - \alpha_{t,2}\right) \cdots \left(1 - \alpha_{t,j-1}\right) \]

\[ + \epsilon_{t,j} \left(1 - \alpha_{t,2}\right) \cdots \left(1 - \alpha_{t,j-1}\right) \]

\[ + \cdots \]

\[ \epsilon_{t,j} \left(1 - \alpha_{t,j-1}\right) \]

\[ + \epsilon_{t,j} \]
3. Moment restrictions for the equation in (generalised) differences

3.1. The standard procedure in linear equations

Anderson and Hsiao (1982) have devised the classic strategy for estimating linear dynamic panel equations with fixed effects. Consider, for example, a first order autoregressive equation with fixed effects:

\[ y_{i,t} = \alpha y_{i,t-1} + \mu_t + \epsilon_{i,t} \]  

(5)

The fixed effect is obviously correlated with the endogenous variable. Transforming the equation by taking first differences eliminates the fixed effect:

\[ \Delta y_{i,t} = \alpha \Delta y_{i,t-1} + \Delta \epsilon_{i,t} \]

The transformed error term is correlated with the transformed regressor. This can be accommodated by using an instrument variable procedure. Anderson and Hsiao propose to use either the lagged first difference, or lagged levels as instruments. Using the second and further lags of the level as instruments for the differenced equation makes use of the following moment restrictions:

\[ E(\epsilon_{i,t-2} \cdot \Delta \epsilon_{i,t}) = 0 \]

\[ E(\epsilon_{i,t-3} \cdot \Delta \epsilon_{i,t}) = 0 \]

\[ \vdots \]

Obviously, using adjacent lagged levels as instruments makes the use of lagged differences redundant. If the moment restrictions hold, then the lagged levels are valid instruments, because they are correlated with the regressor variable. The suggestion of Anderson and Hsiao was further developed by Holtz-Eakin, Newey and Rosen (1988) and Arellano and Bond (1991), who propose to use a more efficient GMM estimator that uses all available moment restrictions optimally, instead of the simple IV method. Formally, the moment equations are written as a system, in order to be able to use varying number of instruments, according to availability. The instruments are weighted optimally using the Hansen (1982) two-stage procedure.

3.2. A transformation made to measure

In this section, we want to adjust the classic estimation strategy for the use with regime dependent error correction models. We start by looking at the first difference of \( y_{i,t} \):

\[ \Delta y_{i,t} = (1 - \alpha_{i,t-1}) y_{i,t-1} + \alpha_{i,t-1} \mu_t + \epsilon_{i,t} - \left( (1 - \alpha_{i,t-2}) y_{i,t-2} + \alpha_{i,t-2} \mu_t + \epsilon_{i,t-1} \right) \]
\[
\begin{align*}
&= (1 - \alpha_{i,t-1}) \Delta y_{i,t-1} + (\alpha_{i,t-2} - \alpha_{i,t-1}) (y_{i,t-2} - \mu_i) + \Delta \varepsilon_{i,t} \\
&= \begin{cases} 
(1 - \alpha_{i,t-1}) \Delta y_{i,t-1} + \Delta \varepsilon_{i,t} & \text{for } \alpha_{i,t-1} = \alpha_{i,t-2} \\
(1 - \alpha_{i,t-1}) \Delta y_{i,t-1} + (\alpha_{i,t-2} - \alpha_{i,t-1}) (y_{i,t-2} - \mu_i) + \Delta \varepsilon_{i,t} & \text{for } \alpha_{i,t-1} \neq \alpha_{i,t-2}
\end{cases}
\end{align*}
\]

We see that there are two consequences of taking first differences of observations that belong to different regimes, as defined by the regime indicator in the preceding period. First, the equation will contain the product of a lagged level term and the differences of regime specific coefficients. This could be accommodated using interaction terms. Second, there is a latent term \(- (\alpha_{i,t-2} - \alpha_{i,t-1}) \mu_i\) that is correlated with all lags of \(y_{i,t}\) and therefore precludes the use of lagged levels in moment equations.

One might think of the following strategy: Differences are only formed for observations with \(r_{i,t-2} = r_{i,t-1}\). On the basis of information from differences with \(r_{i,t-2} = r_{i,t-1} = 0\) we may estimate \(\alpha_0\), and with observations fulfilling \(r_{i,t-2} = r_{i,t-1} = 1\) we can infer on \(\alpha_1\). This idea will not work for two reasons. First, we not only lose the first observation for each individual, but one for every regime switch. Second, the transformed residual \(\Delta \varepsilon_{i,t}\) has an expectation different from zero in the two groups of observations formed by \(r_{i,t-2} = r_{i,t-1} = 0\) and \(r_{i,t-2} = r_{i,t-1} = 1\). This is because \(r_{i,t-1}\) and \(\varepsilon_{i,t-1}\) are correlated. The expectation \(\mathbb{E}(\varepsilon_{i,t-1} | r_{i,t-1} = 0)\) is not equal to zero, and neither is \(\mathbb{E}(\varepsilon_{i,t-1} | r_{i,t-1} = 1)\). Selecting residuals according to regimes will lead to biased estimators.

We want to discuss a strategy that, while following the basic approach of Anderson and Hsiao, circumvents the two drawbacks mentioned above and can be realised using the GMM methodology developed by Arellano and Bond and Holtz-Eakin et al. It consists of the following two principles:

- Let \(q\) be the maximum lag \(k\) for which there is correlation between \(\alpha_{i,t}\) and \(\varepsilon_{i,t-k}\), as a consequence of the MA-structure of the variable \(z_{i,t}\) underlying the regime indicator.

  Then the observation must be transformed subtracting past observations of the same regime with a lag of at least \(2 + q\).

- If an observation is not matched by a \(2 + q\) -lag in the same regime, it may be transformed using any other lag \(l \geq q + 2\).

First, we need to derive the \(l'th\) difference. Similar to eq. (6) we get:
Consider a difference of two observations, \( y_{i,t} - y_{i,t-1} \) characterised by the same regime, such that \( r_{i,t-1} = r_{i,t-1-1} \in \{0,1\} \):

\[
y_{i,t} - y_{i,t-1} = (1-\alpha_{i,t-1}) \left( y_{i,t-1} - y_{i,t-1-1} \right) + \varepsilon_{i,t} - \varepsilon_{i,t-1}
\]

When does the expectation of the residual term, \( \varepsilon_{i,t} - \varepsilon_{i,t-1} \), become zero? It is sufficient that \( \varepsilon_{i,t} \) and \( \varepsilon_{i,t-1} \) are both uncorrelated with the conditioning variables, \( r_{i,t-1} \) and \( r_{i,t-1-1} \). By assumption, \( \varepsilon_{i,t} \) is uncorrelated with \( r_{i,t-1} \) and \( r_{i,t-1-1} \). The same is true with respect to \( \varepsilon_{i,t-1} \) and \( r_{i,t-1} \). Therefore, by choosing \( l \), we only have to make sure that \( \varepsilon_{i,t-1} \) and \( r_{i,t-1} \) are uncorrelated. This will never be possible for \( l=1 \), as we have seen above. However, if \( r_{i,t} \) is uncorrelated with the lag of \( \varepsilon_{i,t} \), then \( l=2 \) is a lag length that gives us:

\[
E \left( \varepsilon_{i,t} - \varepsilon_{i,t-1} \mid r_{i,t-1} = r_{i,t-1-1} = 0 \right) = E \left( \varepsilon_{i,t} - \varepsilon_{i,t-1} \mid r_{i,t-1} = r_{i,t-1-1} = 1 \right) = 0
\]

More generally, if there is correlation between \( \alpha_{i,t} \) and \( \varepsilon_{i,t-k} \) up to the order \( k = q \), the difference that guarantees the equality of conditional expectations will have to be of order \( 2+q \). Furthermore, we are not restricted to using only differences of the order that is "just right", i.e. \( 2+q \). Any other difference of order \( l \geq 2+q \) will fulfil the equality of conditional expectations just as well. Therefore we may take the most proximate difference between two observations of the same regime with \( l \geq 2+q \). With respect to the admissibility and validity of instruments, the rules of the classic approach apply:

**Proposition 2:** Assume the expected value of \( \varepsilon_{i,t} \) to be invariant with respect to all past values \( \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots \), \( r_{i,t-1}, r_{i,t-2}, \ldots \), as well as to \( \mu_i \) and \( y_{i,0} \). Furthermore, the memory of the process driving \( r_{i,t} \) is of finite length \( q \) with respect to \( \varepsilon_{i,t} \), such that the expected value of \( \varepsilon_{i,t} \) is also invariant with respect to the future values \( r_{i,t+q+1}, r_{i,t+q+2}, \ldots \), that is:

\[
E \left( \varepsilon_{i,t} \mid \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots, r_{i,t-1}, r_{i,t-2}, \ldots, r_{i,t+q+1}, r_{i,t+q+2}, \ldots, \mu_i, y_{i,0} \right) = 0
\]

Then the lags \( y_{i,t-1} \) and all earlier lags are valid instruments for the equations transformed by taking the \( l \)th difference, with \( l \geq 2+q \):
and so forth, for all \( c \in \{0,1\} \)

**Proof:** The proposition follows from the law of iterated expectations:

\[
E\left( (\varepsilon_{i,t} - \varepsilon_{i,t-1}) y_{i,t-1} \mid r_{i,t-1} = c \right) = 0
\]

\[
E\left( (\varepsilon_{i,t} - \varepsilon_{i,t-2}) y_{i,t-2} \mid r_{i,t-1} = c \right) = 0
\]

\[
\vdots
\]

\[
E\left( (\varepsilon_{i,t} - \varepsilon_{i,t+q}) y_{i,t+q} \mid r_{i,t} = c \right) = 0
\]

because the conditional expectation within the brackets is equal to zero for \( l \geq 2 + q \). The assumptions above makes sure that the expected values of \( \varepsilon_{i,t} \) and \( \varepsilon_{i,t-1} \) do not depend on values \( y_{i,j-1} \) of the level term. Therefore the additional conditioning on \( y_{i,j-1} \) will not change the conditional expectation.

The discussion shows that it is an important identifying assumption for the process that drives the regime indicator to have finite memory with respect to innovations \( \varepsilon_{i,t} \). If \( r_{i,t} \) is correlated with all past values of \( \varepsilon_{i,t} \), the conditional expectation of the transformed error term resulting from a difference of two observations from the same regime will not disappear. In the context of the underlying estimation problem, this assumption is justified, because following a shock, the financing structure is restored to its target value in finite time.

### 3.3 A test on the validity of the transformation

Thus, in order to use this set of moment conditions for estimation, we need to decide on the length of the memory of the process driving the regime with respect to \( \varepsilon_{i,t} \). This is difficult to do on an a priori basis. There are two simple solutions. The first is to use the Sargan-Hansen test as a test on the appropriateness of the transformation. This is straightforward, as the Sargan-Hansen test is a test on the validity of the moment conditions (9). The drawback is, that the Sargan-Hansen test is generally used as an omnibus test of the specification, including the choice of the instruments. If we employ the Sargan-Hansen test as a means of finding the correct lag length, then our further estimation will be conditional on the test statistic being insignificant. For further purposes, the test is spent.

Alternatively, we may base a test on the fact that the expected value of the residual will not disappear if the lag length chosen is too short. In that case, as we have seen, the choice of observations belonging to one regime or the other will select positive or negative outcomes of \( \varepsilon_{i,t} \), because of the correlation between the regime variable and the error component \( \varepsilon_{i,t} \). If we enter *regime dummies* into our specification, they will be estimated as positive or negative
quantities according to the direction of selectivity, although they should be zero according to the basic specification. Furthermore, we know how these estimates for regime constants are distributed under the null of a correct specification. Using a GMM estimator, they are asymptotically normal, with mean zero and their standard deviation is given by the standard deviation of the coefficient. In effect therefore, the t-value on these coefficients is a valid test statistic.

It may be argued that this test ignores the possibility that the regime specific constants truly belong into the equation. Consider a trend in the term in the brackets of equation (1), that makes the target level of \( y_{i,t} \) change over time:

\[
\Delta y_{i,t} = -\alpha_{i,t-1}(y_{i,t-1} - \kappa t - \mu_t) + \varepsilon_{i,t}
\]  

(10)

Solving for \( y_{i,t} \), we get:

\[
y_{i,t} = (1 - \alpha_{i,t-1}) y_{i,t-1} + \alpha_{i,t-1} \kappa t + \alpha_{i,t-1} \mu_t + \varepsilon_{i,t}
\]

After transforming the equation by subtracting an observation belonging into the same regime, lagged \( l \) periods, we have

\[
y_{i,t} - y_{i,t-l} = (1 - \alpha_{i,t-1})(y_{i,t-1} - y_{i,t-l-1}) + \alpha_{i,t-1} \kappa l + (\varepsilon_{i,t} - \varepsilon_{i,t-l})
\]  

(11)

A regime specific constant may thus be the result of a trending target variable. However, in this case they should be proportional to each other, with a factor of proportionality given by 1 minus the regime specific coefficient on the lagged dependent variable. It is relatively straightforward to test this restriction using the delta method. More generally, they should not be of different sign, as it will be the case if the coefficient on the regime dummy collects the residuals selected for their high or low value.

4. Moment restrictions for the equation in levels

Arellano and Bover (1995) and Blundell and Bond (1998) have proposed to use an additional set of moment conditions that can help identifications of coefficients in autoregressive equations in cases where the level instruments are weak, e.g. when the coefficient of the lagged dependent variable is in the neighbourhood of 1. Under certain conditions, the following moment condition can be used for the autoregressive equation (5):

\[
E\left[\Delta y_{i,t-1}(\mu_t + \varepsilon_{i,t})\right] = 0
\]  

(12)

If \( \varepsilon_{i,t} \) is serially uncorrelated, it is sufficient that \( y_{i,t} \) is mean stationary and displays a constant correlation with \( \mu_t \) for the moment equation to hold. Blundell and Bond (1998) have shown that this implies a relatively weak requirement on the initial conditions: the deviation
of the initial condition from the stationary level needs to be uncorrelated with the stationary level itself.

In our case, the expected value of \( \Delta y_{i,t-1} \), conditional on \( r_{i,t-1} \), will not be equal to zero, as \( \alpha_{i,t-1} \) is correlated with \( \epsilon_{i,t-1} \) and possibly also with some of its lags. The latent variable \( \mu_i \) generically also has an expected value not equal to zero, so that the moment equation (9) does not hold conditional on \( r_{i,t-1} \) taking a certain value. However, under certain conditions we can still use first differences as instruments for levels if we accept that a possible regression constant will not be identified.

We define

\[
\mu_i = \mu^e + \mu^*, \quad \text{with} \quad \mu^e = \mathbb{E}_i \mu_i
\]

The parameter \( \mu^e \) is the expected value over all individuals \( i \), and \( \mu^* \) is the individual deviation from this expectation. Naturally, \( \mathbb{E}_i \mu^* = 0 \). We may rewrite equation (2):

\[
y_{i,t} = (1 - \alpha_{i,t-1}) y_{i,t-1} + \alpha_{i,t-1}\mu^e + \alpha_{i,t-1}\mu^* + \epsilon_{i,t}
\]

(13)

Note that this equation contains a new, regime specific shift term. In estimation, this term can be taken into account using group dummies for regimes \( r_{i,t-1} \). A possible regression constant would be conflated with the group specific terms and is not separately identified.

Investigating under what conditions \( \Delta y_{i,t-1} \) is a valid instrument for the use in equation (13), we arrive at

**Proposition 3**: The conditions a) to c) below are sufficient for \( \Delta y_{i,t-1} \) to be an instrument in equation (14), irrespective of \( r_{i,t-1} \):

\[
\mathbb{E}\left[ \Delta y_{i,t-1} \left( \alpha_{i,t-1} \mu^* + \epsilon_{i,t} \right) \bigg| r_{i,t-1} \right] = 0 \quad (14)
\]

a) The expected value of \( \epsilon_{i,t} \) is invariant with respect to all past values \( \epsilon_{i,t-1}, \epsilon_{i,t-2}, \ldots, r_{i,t-1}, r_{i,t-2}, \ldots \), as well as to \( \mu_i \) and \( y_{i,0} \), i.e.:

\[
\mathbb{E}\left( \epsilon_{i,t} \big| \epsilon_{i,t-1}, \epsilon_{i,t-2}, \ldots, r_{i,t-1}, r_{i,t-2}, \ldots, \mu_i, y_{i,0} \right) = 0 \quad (15)
\]

b) The expected value of \( \mu_i \) is invariant with respect to the entire history of \( \epsilon_{i,t} \) and \( r_{i,t-1} \), and also invariant with respect to the initial deviation, \( y_{i,0} - \mu_i \), i.e.

\[
\mathbb{E}\left( \mu_i^* \big| \epsilon_{i,t}, \{ r_{i,t} \}, \left( y_{i,0} - \mu_i \right) \right) = 0 \quad (16)
\]

**Proof**: By plugging in the backward solution (4) into equation (1) we obtain:
The expected value of the product of $\Delta y_{it-1}$ and the latent term $\alpha_{it-1}^\ast + \varepsilon_{it'}$ needs to be zero, irrespective of whether $r_{it-1} = 0$ or $r_{it-1} = 1$. The two conditions above make sure that this is the case. The instrument is valid because the term that captures the effect of past innovations, $-\alpha_{it-2} \prod_{l=1}^{t-2} (1-\alpha_{i,kl}) \varepsilon_{it'}$, is highly correlated with the corresponding term in the level of the regressor variable $y_{it-1}, \sum_{l=1}^{t-2} \prod_{k=l}^{t-2} (1-\alpha_{i,kl}) \varepsilon_{i,t'}$, and because there is a common dependence on the initial condition $(y_{i,0} - \mu)$. 

It is natural that we have to impose conditions on $\mu$, now that we leave it in the equation instead of differencing it out. The invariance of expected $\mu$ with respect to $\{\varepsilon_{it'}\}$ is rather unproblematic, this accords well with the basic structure of the error component model. According to the assumptions we made with respect to the process that governs $r_{it}$, the irrelevance of the regime process is given as well. However, this condition is not innocuous: it is well conceivable that a real-world data generating process for $r_{it}$ may contain a fixed effect correlated with $\mu$. This would invalidate the moment equation (12), whereas the differencing approach could still be used. Lastly, the necessity of having an expected value of $\mu$ that is independent of the initial deviation, $(y_{i,0} - \mu)$ was also found by Blundell and Bond (1998) when investigating the use of moment equation for levels in a linear context. The condition is not innocuous either, it excludes an initial condition such as $y_{i,0} = 0$. We can replace it by the requirement that the process has been running for a "very long" time, as the first term inside the bracket of equation (17) will disappear asymptotically.

It is interesting to compare the conditions for Propositions 2 and 3. Both require that the expected value of $\varepsilon_{it'}$ be invariant with respect to all past values $\varepsilon_{it-1}, \varepsilon_{it-2}, \ldots, r_{it-1}, r_{it-2}, \ldots$, as well as to $\mu$ and $y_{i,0}$. Proposition 2 needs as an additional identifying assumption that the memory of $r_{it}$ be finite with respect to $\varepsilon_{it'}$. This excludes, for example, an autoregressive equation for $z_{it'}$ in the process driving the regime variable. Furthermore, for consistent estimation we need to find the length $q$ of this memory. The level estimator does not need this second set of assumptions. Instead, we need the expected value of the individual effect $\mu$ to be unrelated to the rest of the process, including the initial deviation $(y_{i,0} - \mu)$. If this assumption is fulfilled, the moment conditions will not depend at all on the memory of the process driving the regime indicator.
5. Implementing and simulating the two estimators

5.1. Setting up the simulation

Both sets of moment conditions have been used separately for estimation. The respective requirements are rather different, and the two sets of moment conditions do not identify the same set of parameters. Therefore it is not straightforward to combine them into a system estimator that uses both sets of moment conditions simultaneously, in analogy to what was proposed by Arellano and Bover (1995) and Blundell and Bond (1998) for linear models. In our opinion, the two sets of moment conditions should rather be used as separate pieces of information giving partly independent evidence.

Both estimators are implemented by first calculating the transformed observations and the instruments, and then adapting and using the routines supplied with the DPD-module for Ox written by Doornik, Arellano and Bond to perform the GMM estimates and the tests. Technically, a new class in Ox is created that derives from DPD (which in turn derives from the Modelbase class and the Database class). When simulating the processes, the following parameters were used:

**Regime dependent error correction process:**
\[
\alpha_0 = 0.7, \quad \alpha_1 = 0.2, \quad \text{so that } 1 - \alpha_0 = 0.3 \quad \text{and } 1 - \alpha_1 = 0.8,
\]
\[
\varepsilon_{i,t} \sim N(0,1), \quad \mu_t \sim N(1,1), \quad \varepsilon_{i,t} \quad \text{and} \quad \mu_t \quad \text{are independent.}
\]

**Regime indicator process:**
\[
E_{\eta_{i,t}} = 0.8, \quad \text{with } \eta_{i,t} \text{ being calculated as a weighted sum of } \varepsilon_{i,t} \quad \text{and an independent Gaussian variable. With respect to the MA structure, we experiment with a MA(0) and a MA(1) with equal weights of 0.5. The variance of the MA is normalised to 1. Note that the correlation of } \eta_{i,t} \quad \text{and} \quad \varepsilon_{i,t} \quad \text{is fairly high, and the larger of the two autoregressive coefficients is not far from 1.}
\]

**Simulated panel structure:**
The panel is unbalanced, with individuals carrying either 7, 8 or 9 observations, 1,000 of each type, that is 3,000 in total. For each individual, the process is simulated for 100 periods, and only the last 7, 8 or 9 observations are used for estimation.

5.2. Generalised difference estimation

When implementing the difference estimator, we use the moment conditions in a specific way that greatly facilitates the calculation of moments. Proposition 2 requires us to calculate the \( l \)'th difference of every observation, with \( l \geq q + 2 \) and differences being taken using only

---

4 Ox is an object-oriented matrix programming language. For a complete description see Doornik (2001).
observations in the same regime. Then we may use the levels lagged $l+1, l+2, \ldots$ as instruments. It seems that this requires us to make the set of instruments for a specific observation dependent on whether or not there are two observations in the same regime within a specific time distance. By taking the earlier of the two observations as a point of reference $y_{i,t}$ and assigning to it the nearest lead $y_{i,t}$ of the same regime with $l \geq 2 + q$, the definition of suitable instruments is straightforward. We can uniformly use lags $y_{i,t-1}, y_{i,t-2}$ and earlier to form the moments

\[
y_{i,t-1} (\hat{e}_{i,t-1} - \hat{e}_{i,t}) r_{i,t-1} = r_{i,t-1}
y_{i,t-2} (\hat{e}_{i,t-1} - \hat{e}_{i,t}) r_{i,t-1} = r_{i,t-1}
\]

with $\hat{e}_{i,t-1} - \hat{e}_{i,t}$ the differenced residual that would result from a given parametrisation of the adjustment equation.

In performing the simulations it turns out that the instruments are much more informative (the estimates being more precise) if the instruments are defined as separate variables, according to their respective regime. That is: For the purpose of instrumentation, the lags of $y_{i,t}$ are stored as different variables, depending on the regime they are in. If the observation is in regime 1, the respective variable for regime 2 is set to zero, and vice versa. The reason for the enhanced precision is probably that the correlation between the regressor variables and the lagged levels is regime dependent, so that the instruments are more informative if they are sorted according to regimes. In the current set of estimations, we have calculated moments only with lagged level variables that fall into the same regime.

5.3. Level estimation

The level estimator is implemented by specifying an equation that contains two sets of dummies, one for $r_{i,t-1} = 0$ and another for $r_{i,t-1} = 1$. Note that these group dummies are not contained in the basic equation (1). They have to be included in order to capture the term $\alpha_{i,t-1} \mu^t$ in equation (13). Not including them would result in a misspecification bias. In order to have more informative instruments, the first differences were saved in separate variables, depending on whether they are differences between two regime 0 variables or two regime 1 variables. Only those lagged differences were used that fall into the same regime as the current observation.

5.4. Simulation results

Column (1) in Table 1 gives the results of a simulation for the case in which there is no autocorrelation in the regime variable, i.e. the variable $z_{i,t}$ in equation (3) follows an MA(0) process. The correct transformation is chosen, with a lead of at least 2 for the formation of differ-
ences. Column (2) does the same for an MA(1) in the process of the variable that drives the regime indicator. Again the correct transformation is chosen, which is a lead of at least 3. Column (3) shows a set of estimates where the process is the same as in column (2), i.e. the regime indicator is driven by a MA(1) process. However, a wrong lead (lead >= 2) is chosen for the transformation. Finally, column (4) and (5) show the results for a level estimation.

The estimation routines in the first columns are consistent, and the estimated standard deviations are in line with the realised RMSE. The t-tests reject the true parameters with a frequency that is not far away from 5%. The Sargan-Hansen test rejects the specification with a frequency that is near the theoretically expected value of 5%. Of course, the standard deviations in column (3) are larger, because the additional lead leads to a loss of observations.

Column (3) does "the wrong thing" – the generalised difference estimator is used with a lead of 2 instead of 3. This shows up in a strong bias of the group specific constant, and a smaller bias of the coefficient $\alpha_i$. There is almost no bias in the estimation of $\alpha_0$. Graph 1 shows the time series of the t-tests on $\alpha_i$ the true value of 0.2, for both the correct and the incorrect lead.

The Sargan-Hansen test is rather sensitive to the violation of the moment condition implied by the inappropriate transformation: the specification is rejected in three out of four cases. Using the t-values of the regime specific constants as a test on the appropriateness of the transformation would lead to a rejection in every single of the 1000 runs. As the estimated coefficients are of opposite sign, they cannot be caused by trending target values.

Column (4) uses the level estimator on a process where there is no autocorrelation in the process driving the regime. The estimator is unbiased, and the test statistics turn out to be correct. The standard deviation of the first coefficient, $1 - \alpha_i$, is clearly higher than in column (1), where the same process is estimated using the generalised difference estimator. On the other hand, the second coefficient, $1 - \alpha_i$, is estimated more precisely.

Column (5) does the same level estimation on the outcome of a process where the regime variable is driven by an MA(1) process. The prediction that the level estimation is unaffected by the memory of the regime process bears out. Actually, the estimates are somewhat more precise compared to those in Column (4), which may be the due to the regime variable being more pervasive. In both Column (4) and Column(5), the regime dummy takes on the expected values of $\alpha_{i,\epsilon,\mu^t}$, a term that is introduced into equation (13) by splitting up the firm fixed effect into its expectation and a deviation uncorrelated with the shocks in the other processes.

5.5. Comparing the two methods

We have presented two different ways of estimating an error correction model with time varying persistence, both within a GMM framework, albeit with a different set of moment conditions. The first technique relies on transformation using generalised differences, with a
lead that is long enough to overcome the memory in the process driving the regime indicator for the $E_{t,i}$-shocks. The second method leaves the equation untransformed, and past differences are used as instruments. We employ regime dummies to catch up and neutralise the time varying non-zero expected value of the residual process. The memory of the regime process is irrelevant for this technique. However, we have to assume the firm specific threshold to be independent of the shock parameters of the other relevant processes. Whereas the generalised difference estimator is good at estimating small coefficients $(1 - \alpha_k)$, the level estimator is more precise with regard to larger coefficients. This is not really surprising: It has been noted by Blundell and Bond (1998) that if the autoregressive parameter approaches 1, lagged levels will be weak instruments when used on a differenced equation, and they propose to use the moment conditions (12) to overcome this problem.

Ultimately, we want to find out whether the speed of adjustment really varies according to the regime. This amounts to testing whether $\alpha_0 - \alpha_1$ is different from zero. The more precise the estimates are, the better we can distinguish $\alpha_0$ and $\alpha_1$. Depending on the size of the coefficients, the two estimators have their comparative advantages, and both of them need specific assumptions that may or may not be fulfilled. Therefore it seems to be wise to use them side by side and hope for a consistent picture to emerge.

6. Taking the estimator to the data

In order to explore the capacity of our estimation method to distinguish between regimes with respect to the dynamics of capital stock accumulation, we estimated a very simple error correction model with one lag, ECM(1), for the real capital stock of the firm. Although quite suggestive, the results presented here are very preliminary, and they have not been checked for robustness. Therefore, at this time, we do not want to take a firm stand in matters of substance.

The equation is derived in the Appendix to this paper, its general form is

$$\Delta \log K_{i,t} = b_0 \Delta \log S_{i,t} + c_0 \Delta \log UC_{i,t} + s \cdot \{ \log K_{i,t-1} + k_1 \log S_{i,t-1} + k_2 \log UC_{i,t-1} \} + u_{i,t}. \quad (18)$$

Here, $K_{i,t}$ is the real capital stock at the end of the period, and its log change is a proxy for the percentage change as a result of current investment. In econometric practice, researchers often put $I_{i,t}/K_{i,t}$ on the left hand side of the accumulation equation. Real sales in period $t$ are denoted by $S_{i,t}$, and the difference of the logs is approximately equal to the growth rate. $UC_{i,t}$ are the user costs of capital; they are permitted to have a firm-specific component. The coefficients are all linear functions of the parameters for an autoregressive investment model with distributed lags (ADL); see the derivation in the Appendix.
The term $s$ in the second line is the error-correction coefficient; it should be negative. The term in curly brackets ought to be zero in the long run for the long-term relationship between capital, user costs and sales to hold. The equation is usually estimated using $\log\left(\frac{K_{it-1}}{S_{it-1}}\right)$ instead of $\log K_{it-1}$ as level term, and we follow this convention. This affects the interpretation of $\log S_{it-1}$: its coefficient will now reflect possible increasing or decreasing returns, and omitting the variable will impose constant returns. The latent term $u_{it}$ is generally supposed to have the following two-way error component structure:

$$u_{it} = \varphi_i + \lambda_t + \zeta_{it},$$

with $\varphi_i$ firm specific, $\lambda_t$ time specific and $\zeta_{it}$ uncorrelated, but possibly heteroscedastic.

We will absorb the user costs into the error term. Depending on whether we allow for time specific shocks $\lambda_t$ by including a full set of time dummies or not, we either assume that user costs are firm specific and constant, or that they are firm specific and subject to shock common to all firms. Of course, the two-way error component structure will also accommodate firm specific technology levels or business cycle dynamics common for the entire manufacturing sector.

| Question 5 [Autumn survey]: Factors influencing investment in 1999-2000 |
|-----------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| In 1999-2000 our investment in western Germany was/is being positively/adversely affected by the following factors: (Please refer to the explanatory notes on the reverse of the accompanying letter.) |
| Sales conditions / expectations               | Very stimulating | No influence | Very stimulating | No influence | Very stimulating | No influence | Very stimulating | No influence |
| Availability / costs of finance               |              |              |              |              |              |              |              |              |
| Earnings expectations                         |              |              |              |              |              |              |              |              |
| Technical development                         |              |              |              |              |              |              |              |              |
| Acceptance of new technologies               |              |              |              |              |              |              |              |              |
| Basic economic policy conditions             |              |              |              |              |              |              |              |              |
| Other factors, namely ...                     |              |              |              |              |              |              |              |              |

The theoretical model predicts that adjustment to a higher target capital stock will be slower when the firm has to overcome financing constraints, compared to a situation where the same gap has to be closed without financing constraints. What we want to test is whether the adjustment dynamics is different depending on the financial regime a firm is in. Depending on
the regime, the coefficients in equation (18) may take two or three different values. It is im-
portant that the regime classification be made on the basis of past information, i.e. with survey
data that arrives prior to the information on the investment in a given year.

The Ifo data set and its information on investment and financing conditions has been dis-
cussed in detail by von Kalckreuth (2004). With respect to whether a firm feels financially
constrained or not, we may use Question 5 of the autumn survey, which asks for the factors
stimulating or limiting investment in the current year and in the coming year. We will use the
answers on financing constraints, as expected for the next year. The financing constraints in-
dicator receives a 1 if financing conditions are regarded as "limiting" or "very limiting", else it
is zero. The current change of the capital stock on the left hand side is using the fixed capital
investment in the year after the autumn survey question was asked. With respect to current
investment, the financing constraints information is predetermined. Similarly, we need infor-
mation on whether the firm is planning to expand or not, as only for expanding firms financ-
ing constraints are predicted to have the retarding dynamic consequences sketched in the In-
troduction. We use the information on the investment purposes from Question IV of the
Spring Survey:

<table>
<thead>
<tr>
<th>Question 4.1 [Spring survey]: Investment purpose in 1999-2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>In 1999 and 2000, our investment was/ will be <strong>predominantly</strong> made to achieve the following purpose:</td>
</tr>
<tr>
<td>(If investment had more than one purpose, please characterise the relative importance by filling in digits 1, 2, ... into the larger boxes)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purpose</th>
<th>2000</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expansion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) not changing the production program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) changing and/or increasing the scope of the production program</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rationalisation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) directed at labour costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) directed at other costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Replacement</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We consider a firm as targeting a higher capital stock if the respondent ticks either "expansion" or "rationalisation directed at labour costs", as the latter will typically be achieved by increasing the capital intensity. We group a firm as not expanding or contracting, if it either ticks "replacement" as main purpose of investment, or "rationalisation directed at other costs". Many firms that report not to have any investment plans in the introductory survey question do not answer to Question 14 at all. In those cases a firm is also classified as not expanding or contracting. Again, this information has to be predetermined with respect to the investment of the current period. We take the information on the purpose of planned investment one year before the spring survey that gives us the information on the current change of the capital
stock. The preparation of the capital stock and sales data is as described by von Kalckreuth (2004).

Table 2 shows the result of estimating the model using the Generalised Difference transformation. In all cases, a lead of 2 was used, as this will leave us with the largest number of observations. A systematic search for the appropriate lead will have to follow. Column (1) and (2) estimate the ECM(1) with the assumption of identical coefficients for all firms at all times. Column (1) specifies the ECM(1) without time dummies, and Column (2) introduces year dummies to accommodate unobserved shocks that affect all firms in the same way. The second block of estimation results, Column (3) and (4), investigates regime specific adjustment, distinguishing just two regimes. A firm in Regime 1 is not financially constrained, whereas a firm in Regime 2 reports to be limited by financing conditions. Again, the equation is estimated both with and without year dummies. The third block finally, Column (5) and Column (6), introduces the pattern of regimes we need to test the theoretical prediction. Regime 1 reflects an expanding firm that is not financially constrained. Regime 2 stands for a firm that is expanding while reporting financing constraints, and Regime 3 is attributed to firm years that are non-expansionary or contracting. Estimation and instrumentation is as described earlier. In all cases, we use lags 1 to 3 of both log capital stock and log real sales as instruments. In the specifications distinguishing between regimes, the instruments are made regime specific as well, so that formally four (in the case of two regimes) or six (in the case of three regimes) continuous-type variables enter the set of instruments. Furthermore, the predetermined regime indicator enters the set of instruments, and also a constant or the set of time dummies. Time dummies are made regime specific in the second and the third block of estimates.

Column (1) shows that the simple ECM(1) model without time-specific shocks and with coefficients that are constrained to be identical in all firm-years is rejected by the Sargan-Hansen test. The model with time dummies included, Column (2), fares better: the Sargan-Hansen is just above the 5% level of significance. The estimation indicates error-correction behaviour, although the speed of adjustment is rather slow: around 15% of a gap is closed each year.

Column (3) and (4) introduce the distinction between constrained and non-constrained firm years. Among the 4889 observations that could be used for estimation, a number of 4059 firm-years are associated with the unconstrained regime, whereas 830 firm years are constrained. The specification without time dummies is rejected by the Sargan-Hansen test. The specification with time dummies is accepted. The estimated speed of adjustment is not very much different in the two regimes, the F-test does not reject their equality. If anything, the speed of adjustment in the constrained regime is somewhat larger (0.227%) than in the unconstrained regime (0.164%). However, the F-test clearly rejects homogeneity with respect to all coefficients involved, the regime specific time dummies excluded.
Our principal hypothesis is addressed only in the third block, Columns (5) and (6). We compare coefficients for unconstrained expansion firm years, constrained expansion, and contraction. With this definition of regime, the Sargan-Hansen test does not reject either the specification with time dummies, or the specification without. Of the available 4037 firm-years of observations for this setting, 2106 are attributed to Regime 1, 382 to Regime 2, and 1549 to Regime 3. For both estimations, the equality of the speed of adjustment coefficients is rejected, as well as homogeneity with respect to all coefficients. In the model without time dummies, the speed of adjustment is estimated at a fairly high 0.326 for unconstrained expanders. Constrained expanders do not seem to exhibit any error correction behaviour at all, their adjustment coefficient is an insignificant 0.036. The contracting firm years are characterised by an adjustment coefficient of -0.155, which is approximately equal to the average depreciation rate and corresponds well to the way the capital stock was calculated using the eternal inventory method. Introducing time dummies leads to a similar pattern, with adjustment coefficients of -0.190, 0.042, and -0.143, respectively. The regime constants are all positive and very small, the constant is significant only for the contraction regime.

7. Outlook

We have developed a method to investigate regime dependent adjustment behaviour using a GMM framework. Our first attempts at testing a hypothesis on the differential speed of adjustment to expansionary shocks of financially constrained and unconstrained firms showed results that are consistent with theory. They also indicate that the method is sensitive enough to detect differences in the adjustment dynamics at real-world sample sizes.

We want to reiterate the caveat in the first paragraph of Section 6. Much remains to be done. In the methodological part of the paper, we have insisted that there are restrictions with respect to the memory of the process underlying the regime variable, and these have to be carefully checked. The level estimator as a second, competing way of estimating regime-specific persistence has still to be implemented with the survey data. A systematic search among different ECM specifications has to be done – although it is not a bad sign that the simplest of all possible investment error correction models seems to work satisfactorily. And additional explanatory variables can be introduced – especially the innovation data of the Ifo survey open up tantalising prospects!

The procedures described here will be useful in other fields, too. Possible applications range from education and labour market issues to household saving and the price setting dynamics of firms.
References


<table>
<thead>
<tr>
<th></th>
<th>(1) MA(0)</th>
<th>(2) MA(1)</th>
<th>(3) MA(1)</th>
<th>(4) MA(0)</th>
<th>(5) MA(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GD, lead = 2</td>
<td>GD, lead = 3</td>
<td>GD, lead = 2</td>
<td>Level est.</td>
<td>Level est.</td>
</tr>
<tr>
<td></td>
<td>correct transform.</td>
<td>correct transform.</td>
<td>inappropr. transform.</td>
<td>no transf.</td>
<td>no transf.</td>
</tr>
<tr>
<td>$1 - \alpha_0$</td>
<td>0.297910</td>
<td>0.298950</td>
<td>0.29687</td>
<td>0.31093</td>
<td>0.30608</td>
</tr>
<tr>
<td>mean estimate</td>
<td>0.029723</td>
<td>0.034735</td>
<td>0.03069</td>
<td>0.04761</td>
<td>0.03394</td>
</tr>
<tr>
<td>mean bias</td>
<td>-0.002087</td>
<td>-0.001051</td>
<td>-0.00313</td>
<td>0.01903</td>
<td>0.00608</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.030628</td>
<td>0.035022</td>
<td>0.03097</td>
<td>0.04902</td>
<td>0.03580</td>
</tr>
<tr>
<td>freq. rejections of true value on 5% conf. level</td>
<td>0.056</td>
<td>0.055</td>
<td>0.050</td>
<td>0.062</td>
<td>0.067</td>
</tr>
<tr>
<td>$1 - \alpha_1$</td>
<td>0.797800</td>
<td>0.796990</td>
<td>0.76235</td>
<td>0.79924</td>
<td>0.79932</td>
</tr>
<tr>
<td>mean estimate</td>
<td>0.034092</td>
<td>0.036122</td>
<td>0.03640</td>
<td>0.02266</td>
<td>0.02002</td>
</tr>
<tr>
<td>mean bias</td>
<td>-0.002198</td>
<td>0.003006</td>
<td>-0.03765</td>
<td>-0.00076</td>
<td>-0.00068</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.033353</td>
<td>0.036095</td>
<td>0.05369</td>
<td>0.02324</td>
<td>0.020722</td>
</tr>
<tr>
<td>freq. rejections of true value on 5% conf. level</td>
<td>0.046</td>
<td>0.046</td>
<td>0.196</td>
<td>0.052</td>
<td>0.063</td>
</tr>
<tr>
<td>$G_0$</td>
<td>0.000050</td>
<td>-0.000171</td>
<td>-0.13143</td>
<td>0.19814</td>
<td>0.19946</td>
</tr>
<tr>
<td>mean estimate</td>
<td>0.016278</td>
<td>0.022707</td>
<td>0.01738</td>
<td>0.04625</td>
<td>0.04439</td>
</tr>
<tr>
<td>mean bias</td>
<td>0.000050</td>
<td>-0.000171</td>
<td>-0.13143</td>
<td>-0.00186</td>
<td>-0.00053</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.016022</td>
<td>0.021840</td>
<td>0.13252</td>
<td>0.04699</td>
<td>0.04493</td>
</tr>
<tr>
<td>freq. rejections of true value on 5% conf. level</td>
<td>0.045</td>
<td>0.044</td>
<td>1.000</td>
<td>0.054</td>
<td>0.062</td>
</tr>
<tr>
<td>$G_1$</td>
<td>-0.001018</td>
<td>-0.001235</td>
<td>0.12708</td>
<td>0.69518</td>
<td>0.69796</td>
</tr>
<tr>
<td>mean estimate</td>
<td>0.016269</td>
<td>0.022694</td>
<td>0.01761</td>
<td>0.03805</td>
<td>0.02859</td>
</tr>
<tr>
<td>mean bias</td>
<td>-0.001018</td>
<td>-0.001235</td>
<td>0.12708</td>
<td>-0.00482</td>
<td>-0.00053</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.016135</td>
<td>0.022161</td>
<td>0.12826</td>
<td>0.03782</td>
<td>0.02931</td>
</tr>
<tr>
<td>freq. rejection of true value on 5% conf. level</td>
<td>0.049</td>
<td>0.049</td>
<td>1.000</td>
<td>0.053</td>
<td>0.063</td>
</tr>
<tr>
<td>freq. rejection by Sargan-Hansen on 5% conf. level</td>
<td>0.049</td>
<td>0.046</td>
<td>0.723</td>
<td>0.061</td>
<td>0.059</td>
</tr>
</tbody>
</table>

1) The theoretically correct value is 0.2, due to the presence of regime specific fixed effects.
2) The theoretically correct value is 0.7, due to the presence of regime specific fixed effects.
Graph 1: t-values for $\alpha_1$, MA(1) regime process, when specifying the correct lead 3

Graph 2: t-values for $\alpha_1$, MA(1) regime process, when specifying incorrectly lead 2
Table 2: Using the Generalised Difference Estimator to compare the regime specific adjustment dynamics in an ECM(1)

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous coefficients</th>
<th>R1: unconstrained growth</th>
<th>R2: constr. growth</th>
<th>R3: contraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>R1</strong></td>
<td>Δlog S_{t,i}</td>
<td>0.19329** (0.02774)</td>
<td>0.01565 (0.04635)</td>
<td>0.21374** (0.04132)</td>
</tr>
<tr>
<td>log(K_{t,i-1}/S_{t,i-1})</td>
<td>-0.25723** (0.03498)</td>
<td>-0.15971** (0.03495)</td>
<td>-0.24842** (0.04613)</td>
<td>-0.16439** (0.04488)</td>
</tr>
<tr>
<td>log S_{t,i-1}</td>
<td>-0.07465* (0.03431)</td>
<td>-0.10603 (0.05807)</td>
<td>-0.08018* (0.04107)</td>
<td>-0.17388* (0.06384)</td>
</tr>
<tr>
<td>Reg. const.</td>
<td>0.0080** (0.00161)</td>
<td>0.00687** (0.00223)</td>
<td>0.00687 (0.00223)</td>
<td>0.00601 (0.00361)</td>
</tr>
<tr>
<td></td>
<td># obs. in reg.</td>
<td>5861</td>
<td>5861</td>
<td>4059</td>
</tr>
</tbody>
</table>

| **R2**        | Δlog S_{t,i}             | 0.14008** (0.07112)      | 0.08170 (0.07775)  | 0.25061** (0.1050) | 0.23877 (0.1393)  |
| log(K_{t,i-1}/S_{t,i-1}) | -0.27195** (0.1054)     | -0.2278* (0.1028)       | 0.03619 (0.1359)  | 0.04278 (0.1072)  |
| log S_{t,i-1} | -0.18019 (0.1107)        | -0.12290 (0.1457)       | 0.08777 (0.1588)  | 0.16283 (0.1845)  |
| Reg. const.   | 0.01337** (0.00401)      | 0.00725 (0.00740)       | 0.00725 (0.00740) | 0.00725 (0.00740) |
|                | # obs. in reg.           | 830                      | 830                | 382             | 382             | 382               |

| **R3**        | Δlog S_{t,i}             | 0.15548** (0.04736)      | 0.049585 (0.05900) | 0.15823** (0.06512) | -0.14297* (0.06544) |
| log(K_{t,i-1}/S_{t,i-1}) | -0.13823** (0.06512)     | -0.12290 (0.1457)       | 0.08777 (0.1588)  | 0.16283 (0.1845)  |
| log S_{t,i-1} | -0.05348 (0.05583)       | -0.05907 (0.08285)      | -0.05907 (0.08285) | -0.05907 (0.08285) |
| Reg. const.   | 0.009477** (0.002800)    | 0.00725 (0.00740)       | 0.00725 (0.00740) | 0.00725 (0.00740) |
|                | # obs. in reg.           | 1549                     | 1549               | 1549            | 1549            | 1549               |

Year dummies:
- none
- full set

Strings of obs. (# firms):
- 1616
- 1616
- 1475
- 1475
- 1312
- 1312

# obs. total:
- 5861
- 5861
- 4889
- 4889
- 4037
- 4037

Tests:
- Sargan-Hansen
  - χ²(39) = 90.3 (p < 0.0005)
  - χ²(39) = 53.4 (p = 0.062)
  - χ²(78) = 103. (p = 0.029)
  - χ²(78) = 69.3 (p = 0.748)
  - χ²(117) = 136 (p = 0.109)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)
  - χ²(117) = 127 (p = 0.245)

- Adjustment speed in R1 and R2 identical:
  - χ²(4) = 8.687 (p = 0.0694)
  - χ²(4) = 27.665 (p < 0.0005)
  - χ²(8) = 16.34 (p = 0.0377)
  - χ²(8) = 18.22 (p = 0.0196)

- All coefficients homogeneous between regimes:
  - χ²(4) = 8.687 (p = 0.0694)
  - χ²(4) = 27.665 (p < 0.0005)
  - χ²(8) = 16.34 (p = 0.0377)
  - χ²(8) = 18.22 (p = 0.0196)

Estimation method: GMM estimation on equations transformed by generalised differences, using a lead of 2. The Sargan-Hansen statistic is a test of overidentifying restrictions proposed by Sargan (1958) and Hansen (1982). The robust standard errors from the second step estimation with a small sample correction based on Windmeijer (2005) are in parentheses; ** significant at the 1% level; * significant at the 5% level. The estimation was done using DPD package version 1.2 on Ox version 3.30 and additional, user written software.
Appendix: Deriving the error correction specification

This specification was invented by Charles Bean (1981), and it was introduced to the micro investment literature by Bond, Elston, Mairesse and Mulkay (2003). The point of departure is the static neoclassical equation for capital demand. Using a generalised CES production function, Eisner and Nadiri (1968) derive the following linear equation from the first-order conditions of profit maximisation:

\[
\log K_t = \theta \log S_t - \sigma \log UC_t + \log h_t, \quad (A1)
\]

with \( \theta = \left( \sigma + \frac{1-\sigma}{\nu} \right) \) and \( h_t = A_t \left( \frac{\sigma-1}{\nu} \right) \cdot (\nu \alpha)^\sigma, \quad (A2) \)

where UC is the user costs of capital, \( A_t \) is productivity and \( \sigma \) and \( \nu \) are the elasticities of substitution and scale respectively. The variable \( h_t \) depends on the time-varying terms \( A_t \).

The elasticity of capital to sales is unity (\( \theta = 1 \)) if the production function has constant returns to scale (\( \nu = 1 \)) or if its elasticity of substitution is unity (\( \sigma = 1 \)), ie in the Cobb-Douglas case.

A log-linear demand equation can also be derived for the case of increasing returns to scale, \( \nu > 1 \). If the firm is rationed on the product market, it will have to solve a cost minimisation problem. Then we have \( \theta = 1/\nu \) in (A2) and \( h_t \) will be a term that depends on relative factor prices and the CES parameters.

We assume that the production possibilities are given by the capital stock at the beginning of the current period. Taking account of installation costs and short-run dynamics in the formation of expectations and adding subscripts for individual companies, we generalise the static capital-demand equation by using distributed lags:

\[
A(L) \log K_{it} = B(L) \log S_{ij} + C(L) \log UC_{it} + \log h_{it},
\]

\( A(L), \ B(L), \ C(L) \) being polynomials in the lag operator, not necessarily of the same degree. With the additional constraints

\[
\frac{B(1)}{A(1)} = 1 \quad \text{and} \quad \frac{C(1)}{A(1)} = 1,
\]

the long-run effects of changes in the level of sales or user costs are the same as in the static model (A1) and (A2). We will use the simplest variant, with a lag of one. This leads to

\[
\log K_{it} = \alpha_t \log K_{i,t-1} + \beta_0 \log S_{i,t} + \beta_1 \log S_{i,t-1} + \gamma_0 \log UC_{i,t} + \gamma_1 \log UC_{i,t-1}
\]

5 The discussion paper version of the latter paper was published 1997.
6 Chirinko and von Kalckreuth (2002) develop this equation explicitly by introducing delivery lags and adaptive expectations. The parameters belonging to the expectation formation mechanism are not separately identified.
We subtract the lagged endogenous variables on both sides, and use the identities
\[ \beta_0 \log S_{i,t-1} + \beta_1 \log S_{i,t-1} = \beta_0 \Delta \log S_{i,t} + (\beta_0 + \beta_1) \log S_{i,t-1} , \] and
\[ \gamma_0 \log UC_{i,t} + \gamma_1 \log UC_{i,t-1} = \gamma_0 \Delta \log UC_{i,t} + (\gamma_0 + \gamma_1) \log UC_{i,t-1} . \]

This leads to the standard ECM specification for a model with one lags:
\[ \Delta \log K_{i,t} = (1-\alpha_1) \log K_{i,t-1} + \beta_0 \Delta \log S_{i,t} + (\beta_0 + \beta_1) \log S_{i,t-1} + 
+ \gamma_0 \Delta \log UC_{i,t} + (\gamma_0 + \gamma_1) \log UC_{i,t-1} + u_{i,t} . \]

A latent term \( u_{i,t} \) has been added. It is composed of a firm-specific constant \( \varphi_i \) that reflects multiplicative firm-specific productivity terms as well as the depreciation rates, a time-specific shock \( \lambda_t \) equal for all firms, and finally an idiosyncratic transitory shock \( \zeta_{i,t} \). In this quite general specification, the data are allowed to determine the adaptation dynamics. Introducing more lags would shift the level terms further back. As a long-run equilibrium elasticity of the capital stock with regard to sales, one obtains:
\[ \frac{d \log K}{d \log S} \bigg|_{S=S_{i,t-1}=\ldots=S} = \frac{\beta_0 + \beta_1}{1-\alpha_1} , \]
and a similar expression for the user cost sales elasticity, with the \( \gamma_0 \) and \( \gamma_1 \) in the numerator. Note that the numerator and denominator of this expression are given by the coefficients of the level terms. By rewriting the equation, this identity offers a possibility of neatly separating short-run and long-run dynamics:
\[ \Delta \log K_{i,t-1} = (1-\alpha_1) \log K_{i,t-1} + \beta_0 \Delta \log S_{i,t} + \gamma_0 \Delta \log UC_{i,t} + 
+ (1-\alpha_1) \left[ \log K_{i,t-1} - \frac{\beta_0 + \beta_1}{1-\alpha_1} \log S_{i,t-1} - \frac{\gamma_0 + \gamma_1}{1-\alpha_1} \log UC_{i,t-1} \right] + u_{i,t} . \]

The first term in the third line is the error-correction coefficient; it should be negative. The term in curly brackets ought to be zero in the long run for the long-term relationship between capital, user costs and sales to hold. As in the text, the equation can also be estimated using \( \log \left( \frac{K_{i,t-1}}{S_{i,t-1}} \right) \) and \( \log S_{i,t-1} \) instead of \( \log K_{i,t-1} \) and \( \log S_{i,t-1} \) as level terms. The coefficient of \( \log S_{i,t-1} \) will then reflect possible increasing or decreasing returns, and omitting the variable will impose constant returns. In the absence of user cost information, we may use time dummies to catch the effect of changing relative prices.