PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Department for International Development funded programme run by the Centre for Economic Policy Research (CEPR).

This is a PEDL Research Paper which emanates from a PEDL funded project. Any views expressed here are those of the author(s) and not those of the programme nor of the affiliated organizations. Although research disseminated by PEDL may include views on policy, the programme itself takes no institutional policy positions

PEDL Twitter

Investment Dynamics, Unobserved Heterogeneity and Endogenous Investment Switching Regime in Manufacturing

By Samuel V. Mhlanga*

Abstract

This paper investigates the industrial effects of *true* state dependence, the sales-to-capital ratio and unobserved heterogeneity on the rate of investment in plant, machinery and equipment (PME) in Swaziland. A range of fixed and random effects estimators are compared. In all the methods used, the first-order autoregressive – AR(1) – model with unobserved firm-specific effects and the sales-to-capital ratio have insignificant coefficients. Similarly, the impact of unobserved firm-specific characteristics underlying investment decisions is also insignificant. Most interestingly; however, our novel result is that we show how missing investment values reduce the probability of investing under both exogeneity and endogeneity assumptions. Missing investments at time $t-1$ reduce the likelihood of investing at time t by $[-5.56\%, -4.91\%]$ depending on regressor exogeneity or endogeneity assumptions, respectively. By interacting missing investments with labour, a probability of up to 0.55% of capital substitution for labour is estimated. Furthermore, notice that the Generalized Method of Moments (GMM) and multilevel methods naturally assume a single investment regime by default. When an endogenous switching regime model of investment with unobserved separation is estimated, it produces negative but significant results in both regimes. Notably, all the switching regime results are scale-independent, where firm size is defined as the inverse of the previous period capital stock. However, a Wald-test of equation independence across regimes confirms a single investment regime produced by the GMM and multilevel models.

Keywords: Investment dynamics, heterogeneity, switching regime, manufacturing, Swaziland

Journal Classification: L6, C01, D22, C23

* Corresponding author: smhlanga@uniswa.sz, Tel: +268 7611 9516, University of Swaziland, University Planning Centre, Private Bag 4, Swaziland. I owe a great deal of gratitude to my PhD dissertation advisor Professor Neil Rankin for his guidance and support, and to my Examination Committee for insightful suggestions on the chapter version of this paper. I am grateful to the Private Enterprise Development for Low-Income Countries (PEDL) research initiative of the Centre for Economic Policy Research (CEPR) and the Department for International Development (DFID) for the Exploratory Research Grant (Ref. 1670) covering the cost of buyout, data collection, digitising of records, data analysis and travel between the University of Swaziland and the University of Stellenbosch. My gratitude also goes to the Micro Finance Unit in Swaziland for research assistance funding. I am indebted to the Central Statistical Office (CSO) of Swaziland for making the firm-level dataset available for research. All views expressed in this paper are those of the authors and do not necessarily represent those of, and should not be attributed to, the financier of the project, the (CSO), the University of Stellenbosch or the University of Swaziland. I thank seminar participants at the University of Warwick, CEPR in London, ERSA at the University of Cape Town, CSAE at the University Oxford, particularly my discussant, Monica Paganini, for invaluable comments. All remaining errors are mine. © Author 2018.

1. Introduction

The purpose of this paper is to estimate a dynamic structural model of industrial investment in Swaziland for 1994-2003: a period of trade liberalization in the Southern African Customs Union (SACU). This is an interesting period in its own right because of observed micro churning dynamics and industrial reorganization induced by the trade reforms in the sub-region. It is also interesting because it uses an unbalanced firm-level panel data set that has never been used before to identify determinants of industrial investment decisions in one of the peripheral SACU member economies. The achievement of this goal is important for both policymakers and practitioners working in the field to investigate investment patterns in plant, machinery and equipment (PME) in the presence of a high incidence of zero investments.

Typically, the framework of analysis relates the investment rate at time t to its own $t-1$ realizations *aka* structural state dependence, the marginal q and control variables as explanatory regressors. Structural state dependence is a relationship between the current and the probability of future investment. With structural state dependence, the conditional probability of positive investment in capital goods is a function of past capital investments, see Heckman (1981b). One explanation for this offered in the literature is that preferences, prices and constraints that are fundamental to future investment choices can be directly altered. Another explanation is that firms may differ in certain unobserved firm-specific characteristics underlying their propensity to invest in capital goods. If unobserved heterogeneity is correlated over time, and is not controlled for, past investment may appear to be a genuine cause of future investment simply because it is a proxy for persistent unobservables. In a structural model of investment, it is important to distinguish between the two explanations in order to design appropriate industrial policies that promote firm-level investment.

Tobin's assertion that investment is a function of marginal q and that it is also equivalent to the firm's optimal capital accumulation problem with adjustment costs is now widely recognized, see Hayashi (1982), Caballero and Engel (1999), and Cooper and Haltiwanger (2006). The variation in the structural model of investment is therefore explained by the variation in the shadow price of capital, *or* marginal q. Although marginal q is *a priori* appropriate for characterizing the relationship between movements in the shadow price of capital with investment variation, its unobservability makes it only indirectly applicable in empirical work, see Caballero and Leahy (1996).[1] An alternative candidate is the ratio of the firm's stock market value to its capital replacement cost; that is, Tobin's average q. Caballero and Leahy (1996) argue that Tobin's q is potentially a better covariate in investment regression analyses than marginal q in the presence of fixed costs of capital adjustment. They also

[1] One exception is Gala (2015) who abstracts away from the counterfactual capital adjustment cost assumptions to develop a state-space measure of marginal q that is anchored on the joint measurability of the market value of the firm and its underlying state variables.

provide conditions for it to be a sufficient statistic of capital. However, most industrial firms in Swaziland are not traded in the stock exchange and therefore one cannot use the market value of the firm in constructing a proxy for marginal q. Furthermore, as in Nielson and Schiantarelli (2003), the fact that the data set does not distinguish between multi-plant and single-plant firms, it is not clear how firm-level stock valuations need to be used.

Under the same conditions; nonetheless, the ratio of sales-to-capital is a sufficient statistic of investment. Eberly, Rebelo and Vincent (2012) suggest that simultaneous inclusion of both the sales-to-capital and Tobin's q as regressors might constitute informational redundancy, on condition there is no measurement error in q (also see Erickson and Whited, 2000 for a detailed discussion of measurement error in q). In a structural model of investment, Letterie and Pfann (2007) use the sales-to-capital ratio, average profit of capital and the profit rate as proxies for marginal q.

An extension of this framework is provided by Abel and Eberly (1994) who rely on the theory of investment under uncertainty. In this case, non-convexity, a wedge between the procurement and sale price of capital as well as potential investment irreversibility are key ingredients of their exposition. As is typical, investment is a non-decreasing function of the shadow price of installed capital. This permits identification of firms that sort into a high or low investment regime under conditions of *ex ante* known or unknown sample separation (see Nabi, 1989 and Hu and Schiantarelli, 1998).

This paper therefore estimates a structural model of investment that determines the impact of the lagged response, the proxy of marginal q and unobserved heterogeneity in manufacturing in Swaziland. A good understanding of the driving forces of investment dynamics is crucial for designing well-functioning incentives for industrial development. It requires a distinction between *true* state dependence of investment and its spurious form. The presence of state dependence in firm-level investment data means that industrial policy that encourages current investment improves the probability of future investment[2].

The empirical distinction between longitudinal or within-firm dependence induced by previous realizations and the dependence caused by unobserved heterogeneity is important in studies of dynamic panel data (DPD). In such cases, when investment is treated as a continuous dependent variable, methods for solving initial conditions problems are now standard in DPD models in econometrics, see Anderson and Hsiao (1981, 1982), Arellano and Bond (1991), Blundell and Bond (1998) and Bun and Windmeijer (2010). Corresponding methods for handling the initial conditions problem in discrete response settings are less well developed and are scattered all over the literature.

[2] For example, Christiano, Eichenbaum and Evans (2005) predict joint presence of lagged investment effects together with cash-flow and q effects in an investment model. In a study by Eberly *et al*. (2012) based on the same framework, the lagged investment rate variable has a stronger effect on the current investment rate than the effects of q and cash-flow combined.

In the binary case, Heckman (1981a) models the initial dependent variable jointly with its subsequent response while Wooldridge (2005) conditions on the initial response. In Skrondal and Rabe-Hesketh (2014), these pieces are put together in a multilevel modelling setting to handle initial conditions and covariate endogeneity for dynamic models of binary decisions under unobserved heterogeneity.[3] This approach is applied by Drakos and Konstantinou (2013) to a Greek manufacturing panel dataset.

Our empirical strategy implements the Generalized Method of Moments (GMM) approach to estimate the impact of the previous investment response and other covariates on the current level of investment. Explanatory variables include the proxy for marginal q and control variables, in this case the logarithm of employment level. This also allows us to determine the impact of primary input substitutability during episodes of economic reforms and heightened uncertainty. We also use two competing modelling approaches. The first one is a joint model of initial conditions and subsequent response based on the factor modelling approach, see Bock and Lieberman (1970) and Aitkin and Alfo (2003). This approach allows us to distinguish between exogeneity and endogeneity of explanatory variables. The second one models the distribution of the random intercept conditional on initial conditions and covariates. In order to relax the normality assumption of the random intercept, we also use nonparametric methods to estimate the conditional model, see Heckman and Singer (1984) and Rabe-Hesketh *et al*. (2003) for details. We finally extend the GMM and multilevel investigations to endogenous switching regime regressions in order to establish whether or not firms switch between high and low investment regimes, see Maddala (1983), Dutoit (2007), Hu and Schiantarelli (1998), Nielson and Schiantarelli (2003).

Our findings are that true state dependence and unobserved heterogeneity in the structural model of investment for the manufacturing sector during the trade liberalization period have insignificant effects on investment. The results are consistent with firms exercising their option to wait until uncertainty is resolved, leading to significant substitution of capital for labour. Specifically, firms concentrated more on maintaining and repairing existing machinery and equipment rather than investing in new physical capital. This implies a generally high rate of obsolescence in capital assets and therefore low capital productivity. At the same time, the missing values of investment substituted investment for employment by up to 0.55 percent and reduced the likelihood for future investment by 5.56 percent.

Our contribution to the investment body of knowledge lies in three areas. Firstly, the high incidence of missing values of the response variable means that purging fixed effects using first-differences magnifies the gaps in the transformed unbalanced panel. However, the comparative strength of this transform is that longer lags of regressors remain orthogonal to the noise and available as instruments,

[3] The specific Skrondal−Rabe-Hesketh model is designed for the human health sciences applied to children’s wheezing.

see Roodman (2009a). Nonetheless, in order to minimize data losses arising from the first-difference transform, we use instead the Helmert's transformation to implement the forward orthogonal deviations, see Arellano and Bover (1995) and Roodman (2009a). Secondly, the untransformed data structure also means that the Heckman (1981a) and Wooldridge (2005) methods for estimating dynamic random effects models are faced with an insufficient observations problem when estimating state dependence and random-intercept effects. We overcome this hurdle, to our knowledge for the first time in investment analysis, by reverting to novel techniques proposed by Skrondal and Rabe-Hesketh (2014) which do not insist on balanced panel data to efficiently deal with initial conditions and endogenous regressors. Finally, a range of multilevel dynamic random-effects probit model estimators is performed for comparison with the GMM results and also for extensive comparison of results among the random-effects estimators. Like Stewart (2007), we use normal heterogeneity in the joint and conditional models to handle initial conditions and endogeneity problems. In addition, we also use nonparametric maximum likelihood (NPMLE) methods to estimate the random-effects models.

This paper is organized as follows: The next section describes the panel dataset and performs descriptive analyses of industrial investment rates. In Section 3, the shape of the empirical hazard and fixed adjustment costs are investigated. Section 4 discusses econometric estimators and empirical results of the structural model of investment are presented in Section 5. Section 6 discusses the results and Section 7 summarises and concludes the analysis.

2. Data And Descriptive Analysis

This section focuses on the diagnosis of the data set by describing a few features that are suggestive of the relevance of the organizing framework outlined in the introduction. The dataset consists of an unbalanced census panel of manufacturing firms collected by the Central Statistical Office (CSO) in Swaziland for the period 1994-2003. Although this is referred to as a census because the data collection instrument is administered to all respondents in the sector, the response rate falls short of 100 percent. As a result, a total of 227 firms and 1 448 plant−year observations populate the dataset. However, although there is nonresponse by some firms, missing responses from those that contribute significantly to sectoral GDP are followed up until they return the data collection instruments. In the case of (dis)investment variable response, expenditure in and sales of PME are reported either with missing values or with real numbers.

In structural modelling of investment, movement in investment rates is a function of variation in the sufficient statistics of capital identified by Caballero and Leahy (1996) and Letterie and Pfann (2007). The sufficient statistics are capital ratios of cash flow $[CF_t/K_{t-1}]$, sales revenue $[S_t/K_{t-1}]$ and operating profits $[P_t/K_{t-1}]$, all measured in constant values and expressed in natural logarithm. It is now standard to consider such statistics as proxies of the marginal q, see Gilchrist and Himmelberg (1998) and Letterie and Pfann (2007). The unique feature of our dataset relative to other case studies is that it has disaggregated information on expenditure and sales of capital assets and thus these sufficient statistics can be calculated.[4]

Table 1 presents summary statistics for selected variables of interest in the sample. All the variables are mesokurtic; that is, the mean is always greater than the median, except for real capital stock (K_t). Investment rates (I_t/K_{t-1}) and the associated proxies for the shadow price of capital are positively skewed, suggesting a small fraction of larger firms are distributed along the right fat tails. The variability in a typical proxy of marginal q is approximately $1\frac{1}{3}$ times higher than that of the investment rate. The investment rate variation measured by the standard deviation is relatively low at 0.29, with an average investment rate of 0.24 and $(\text{Min}, \text{Max}) = (-0.83,\ 1.58)$.[5] It is striking that the investment rate and all proxies reveal no marked patterns of heterogeneity across firms. That is, the behaviour of each proxy over time is insignificantly different from the orders of magnitude of other proxies. Hence, a choice to use *any* one of the proxies to study the behaviour of investment rates is likely to suffice.

Table 1: Summary Moments of Key Variables

	Key Variables					Proxies of Marginal q		
Statistics	Emp_{t-1}	I_t	K_t	K_t^{-1}	I_t/K_{t-1}	S_t/K_{t-1}	CF_t/K_{t-1}	P_t/K_{t-1}
Mean	3.55	12.46	9.54	0.11	0.24	1.29	1.29	1.18
Median	3.22	12.34	9.61	0.10	0.20	1.23	1.23	1.11
Std Dev	1.54	2.70	1.51	0.03	0.29	0.36	0.36	0.34
Std/Mean	0.43	0.22	0.16	0.27	1.21	0.28	0.28	0.29
Skewness	0.69	−0.10	−0.47	8.29	1.15	8.47	8.49	7.89
kurtosis	3.11	4.08	5.28	124.42	6.72	116.26	116.55	102.48
IQR	1.96	3.69	1.79	0.02	0.33	0.22	0.22	0.20
Observations	1288	533	1267	1267	401	911	911	907

Key: Emp_{t-1} denotes the log of $t-1$ stock of employment, I_t is the log of net investment in plant, machinery and equipment, K_t represents the log of capital stock at time t, S_t is time t log of real sales revenue from firm output, CF_t is the log of cash-flow at time t and P_t refers to time t log of profits.

There are at least two explanations for the patterns observed in Table 1. First, the Swaziland Government initiated a programme of factory-shell construction in the 1990s to promote foreign direct investment in manufacturing. Specifically, the Textile as well as the Clothing and Wearing Apparel industries were the main beneficiaries of the factory-shell programme due to the AGOA arrangements. This had the effect of reducing private sector capital expenditure on building construction in the sector. Thus, the composition of firms' portfolios of capital goods mostly included

[4] Whenever capital retirement is available in datasets in the literature, it includes the scrap value of capital disposals as a result of obsolescence and sale of capital, see in Cooper and Haltiwanger (2006). In Nielsen and Schiantarelli (2003), net investment is defined as expenditure *minus* sales of fixed capital.

[5] Cooper and Haltiwanger (2006) report an average rate of investment of 12.2 and a standard deviation of 33.7 for NT=100,000 covering large plants that were in continual operation during 1972-1988.

machinery and equipment. Second, the low investment level in PME may be a reflection of risk aversion translating into firms' decisions to exercise the option to wait until the uncertainty induced by economic reforms declined to acceptable levels.

Table 2 presents a correlation matrix of investment rates and proxies of marginal q. The first moment is first-order serial correlation of investment and is estimated at 0.61. A relationship between the current investment and its lagged level suggests a potential presence of state dependence. Similarly, corporate financial performance in the manufacturing sector in Swaziland is almost scale-invariant; i.e., the correlation coefficient between all marginal q proxies and the inverse of capital stock, ${K_t}^{-1}$, is at most 0.03.

Table 2: The Correlation Matrix of the Main Variables

	$\frac{I_t}{K_{t-1}}$	$\frac{I_{t-1}}{K_{t-2}}$	$\frac{I_{t-2}}{K_{t-3}}$	$\frac{I_{t-3}}{K_{t-4}}$	$\frac{I_{t-4}}{K_{t-5}}$	${K_t}^{-1}$	Emp_{t-1}	$\frac{S_t}{K_{t-1}}$	$\frac{CF_t}{K_{t-1}}$	$\frac{P_t}{K_{t-1}}$
I_t/K_{t-1}	1.00									
I_{t-1}/K_{t-2}	0.61	1.00								
I_{t-2}/K_{t-3}	0.42	0.52	1.00							
I_{t-3}/K_{t-4}	0.45	0.60	0.66	1.00						
I_{t-4}/K_{t-5}	0.45	0.53	0.62	0.75	1.00					
${K_t}^{-1}$	0.43	0.21	0.04	0.12	0.05	1.00				
Emp_{t-1}	0.40	0.38	0.47	0.62	0.80	0.18	1.00			
S_t/K_{t-1}	0.26	0.71	0.33	0.33	0.33	0.03	0.13	1.00		
CF_t/K_{t-1}	0.26	0.72	0.33	0.33	0.33	0.03	0.13	1.00	1.00	
P_t/K_{t-1}	0.33	0.76	0.33	0.37	0.38	0.02	0.13	0.96	0.97	1.00

Key: Emp_{t-1} denotes the log of $t-1$ stock of employment, I_t is the log of net investment in plant, machinery and equipment, K_t represents the log of capital stock at time t, S_t is time t log of real sales revenue from firm output, CF_t is the log of cash-flow at time t and P_t refers to time t log of profits.

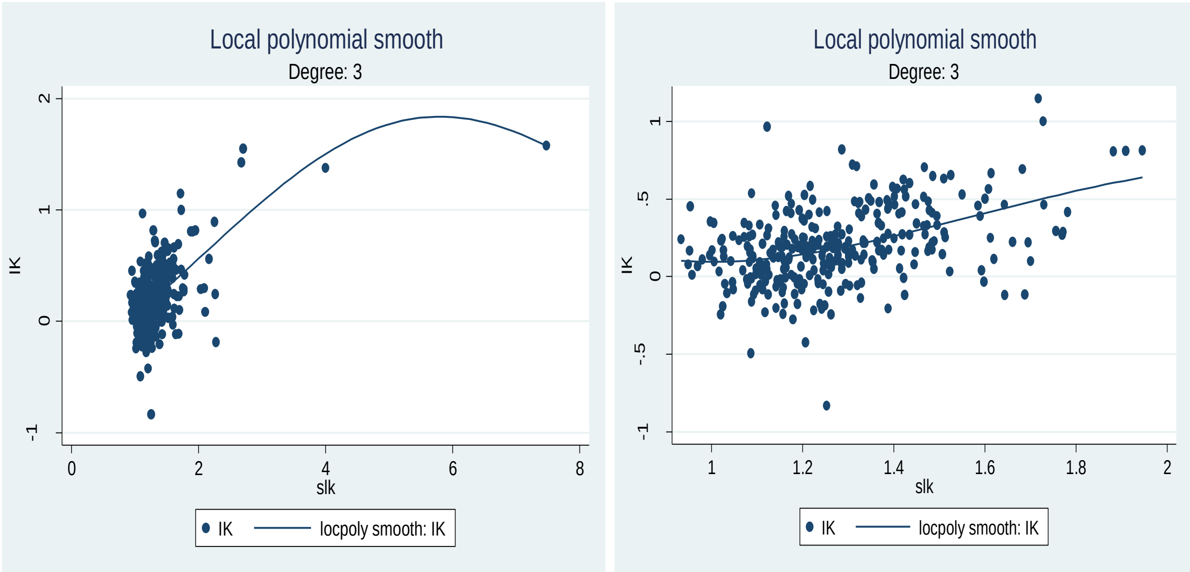
In the correlation matrix, there is low correlation between contemporaneous investment rates and each proxy measure of marginal q. However, the relationship increases significantly to over 0.71 if we look at $t-1$ investment rates and proxies. This suggests that establishments make sales first and then assess the business capital needs before making investments. Thus, there are high investment rates during periods of high sales revenue, high cash flows and high profitability in the sector. As expected, the correlation among marginal q proxies is *at least* 96 percent. From this point forward, our discussion focuses only on the relationship between investment rates and sales revenue as in Letterie and Pfann (2007) for the Dutch case. Similarly, Figure 2 also reports relatively high fourth-order serial correlation in the plant-level investment rate series. This is consistent with the commonly held perception of high autocorrelation of shocks to demand and productivity[6].

A further characterization of patterns of investment (I_t/K_{t-1})-marginal q relation based on S_t/K_{t-1} is graphically presented in Figure 1. In the first panel, a local polynomial smooth of investment rates plotted against the real sales/capital ratio shows a high frequency distribution around an average of

[6] See Cooper and Haltiwanger (2006:614).

1.29 with a standard deviation of 0.28. This panel considers all observations, including outliers. The right panel considers the distribution of plant-year observations for $S_t/K_{t-1} < 2$, where the clustering of observations becomes more sparsely populated. The distribution in this case shows the majority of firms that are consistent with the property that $I_t/K_{t-1} \in [-1,\ 1.2]$.

Figure 1: Investment Rate Relationship with the Sales/Capital Ratio



As in Cooper, Haltiwanger and Power (1999) and Cooper and Haltiwanger (2006), the rest of this chapter defines net investment in terms of real gross expenditure (EXP_{it}) on PME and real sales ($SALES_{it}$) for firm i at time t for the class of capital goods concerned. One striking feature of the expenditure series is that it isolates the cost of maintenance and repairs, permitting a sharper investigation of non-smoothness of (dis)investments. Hence

$$I_t = EXP_t - SALES_t \qquad (1)$$

and

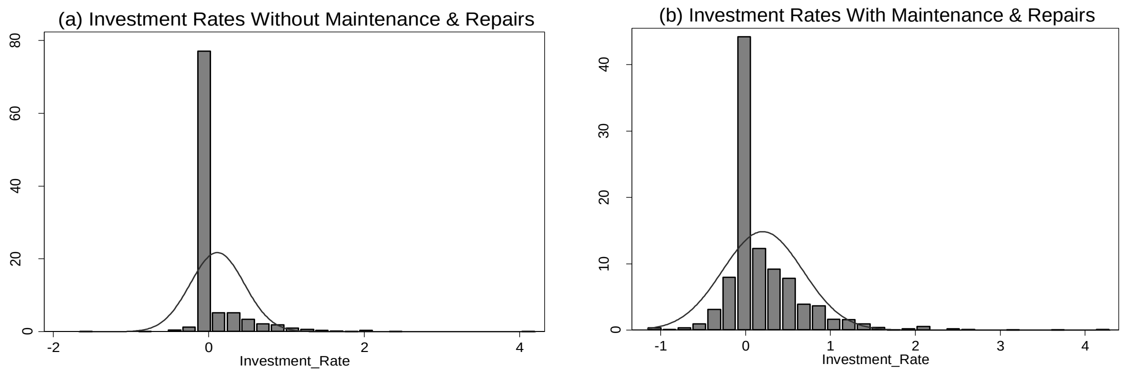
$$K_{t+1} = (1 - \delta_t)K_t + I_t, \qquad (2)$$

which is the perpetual inventory method (PIM) of estimating capital stock, where K_t is the measure of real capital stock, δ_t is the in-use depreciation rate. In Figure 2, the data set is sliced into two non-normal histograms of investment rate with and without maintenance and repairs in panels (a) and (b), respectively. It is characterized by significant mass around zero, fat tails, considerable skewness to the right and high kurtosis.[7] That is, there is a high incidence of zero investments with only a few occasions of lumpy net expenses on capital goods whether or not the cost of maintenance and repairs is included. This exact pattern of investment rate distribution remains unchanged even if the data set is

[7] Standard tests of normality yield strong evidence of skewness and kurtosis at $p < 0.0000$.

sliced to remove outliers as observed in Figure 4.1a, which reduces the observations by 50 percent. Furthermore, Table 4.1 reports a skewness of 1.15 and a kurtosis of 6.72 while the investment rate distribution for the sample of outlying observations reports skewness and kurtosis of 2.11 and 8.48, respectively. The characteristic skewness and kurtosis of the investment rate distribution without the cost of maintenance and repairs remains valid in Figure 2a. These investment distributional patterns have been found in the literature to characterize investment behaviour even at the aggregate level, see Caballero *et al*. (1995), and Doms and Dunne (1998). The pronounced level of skewness, high kurtosis and significant mass around zero in the distribution of investment rates is indicative of the presence of nonconvexities in the capital adjustment technologies. The observed fat tails in Figure 2b suggest the presence of a small fraction of large capital adjustments due to sales and procurement.

Figure 2: Distribution of Investment Rates of PME.



In summary, the analysis thus far provides several lessons. It reveals that there is low propensity to invest in capital goods and that the observed heterogeneity in the rate of investment is just as low. As is typical in the literature, investment inactivity dominates the distribution of investment rates, whether or not maintenance and repairs (M&R) are accounted for. The cross-sectional distribution of investment rate is characterized by skewness and high kurtosis, suggesting the presence of nonconvexities in the capital adjustment costs. Firm-level investment behaviour, including financially unconstrained firms, is also consistent with increased focus on M&R by firms while participating rarely in lumpy investment. These patterns imply the presence of low costs of capital adjustment in manufacturing and a high rate of obsolescence in machinery and equipment needed for use in production. The significantly low level of investment in relation to the observed variation in any one of the chosen sufficient statistics indicates that firm-level revenue is not reinvested in capital goods. Hence, the sales/capital ratio may not have any explanatory power on investment rate changes.

However, the investigation thus far has assumed that investment behaviour of firms is homogenous across size categories. In their study of determinants of African manufacturing investment, Gunning and Mengistae (2001) find that the profit rate of small plants has a significant impact on investment rate while remaining insignificant in the case of large firms. The next section studies the shape of the investment hazard and fixed adjustment costs by firm size category.

3. The Shape of the Hazard and Fixed Adjustment Costs

This section investigates patterns of investments in PME to determine spells of inactivity prior to an investment spike. We follow Kalbefleisch and Prentice (2002) and Cameron and Trivedi (2005) who define the cumulative distribution function representing the probability of a spell length of inactivity as

$$F(t) = Pr(T \leq t)$$

The sample survivor function, $S(t) = Pr(T > t) = 1 - F(t)$, is a step function that decreases by n^{-1} at each observed time t, where n is the number of firms at risk of experiencing an investment spike. It is useful to express the probability of a firm staying in the zone of inaction until time t_j using the nonparametric Kaplan-Meier estimator of the survivor function , $\hat{S}(t)$

$$\hat{S}(t) = \prod_{j|t_j<t} \frac{n_j - d_j}{n_j}$$

where d_j is the number of firms experiencing an investment spike. The Kaplan-Meier estimator, or product limit estimate, calculates the probability of investment inactivity past time t, or the probability of a lumpy investment after time t. This measure precisely aligns with the observed proportion (d_j/n_j) of the n_j firms at risk of experiencing a spike, see Kalbefleisch and Prentice (2002:16).

In order to estimate this model, it is pertinent to define an investment spike and what constitutes the zone of investment inactivity. Economic theory provides no guidance concerning the definition of a lumpy investment episode. However, Cooper *et al.* (1999) use gross investment rate in excess of 20 percent to represent an investment spike. There are some exceptions to this rule. These include Bigsten *et al* (2005) who define a spiky investment as gross investment rate in excess of 10 percent. Studies by Cooper *et al*. (1995) and McClelland (1997) argue and demonstrate that the shape of the hazard rate is robust to any choice of an *ad hoc* definition of a spiky threshold. In this paper, we adopt the definition provided by Cooper *et al.* (1999). Additionally, we define the zone of inaction in terms of investment rate that is bounded as $\frac{I_{it}}{K_{it-1}} \in [-0.049,\ 0.049]$, rather than the standard restriction of $\frac{I_{it}}{K_{it-1}} = 0$.[8]

In spite of definitional modifications, it is possible to ask whether firm-level investment lumpiness is the same across firm sizes during this period. It is of interest to compare the empirical distributions of the survival patterns of large versus small firms' lumpy investment episodes to determine if both samples arose from identical survivor functions. In the left panel of Figure 3, firm scale-independence means the null hypothesis $H_0\colon \text{Survivor}_{\text{Large}} = \text{Survivor}_{\text{Small}}$ is not true, where large firms employ more than 50 workers. The Peto-Peto-Prentice test does not support the null at standard levels of significance.[9] This means the distributions of survival rates for larger (size=1) and smaller (size=0) firms past time *t* are significantly different to each other. It can be concluded that larger firms experience lumpy investments relatively more often than their smaller counterparts.[10] Put differently, the probability of smaller firms staying in the zone of investment inaction is higher than that for larger firms. This suggests that the frequency of investment spikes is scale-dependent in the Swazi manufacturing sector.

Another important area of duration analysis for firm-level investments involves the shape of the hazard estimate. For example, Cooper *et al*. (1999) allow for several characteristics of investment in their machine replacement model to identify three specific patterns of the hazard. *First*, when exogenous shocks to plants' profitability are serially correlated and some additional assumptions hold, the likelihood of capital asset replacement increases with the time since the last replacement. *Second*, adding convex adjustment costs to the autocorrelation assumption ensures the presence of serial correlation in investments and therefore a downward sloping hazard. *Third*, a combination of autocorrelation in exogenous shocks and the absence of adjustment costs produce a flat hazard.

In order to investigate the shape of the hazard in the Swazi data, we first define the probability of experiencing a spike, conditional on remaining in the zone of inaction until time t, as

$$p_{ijt} = pr[T_{ij} = t | T_{ij} \geq t, t - (T_{ij-1} + 1)], \tag{4}$$

where $t - (T_{ij-1} + 1)$ represents the interval since the last spike, while *t* denotes calendar time. We then define discrete time as T_{ij} at which plant *i* exits the state of inactivity to have an investment spike

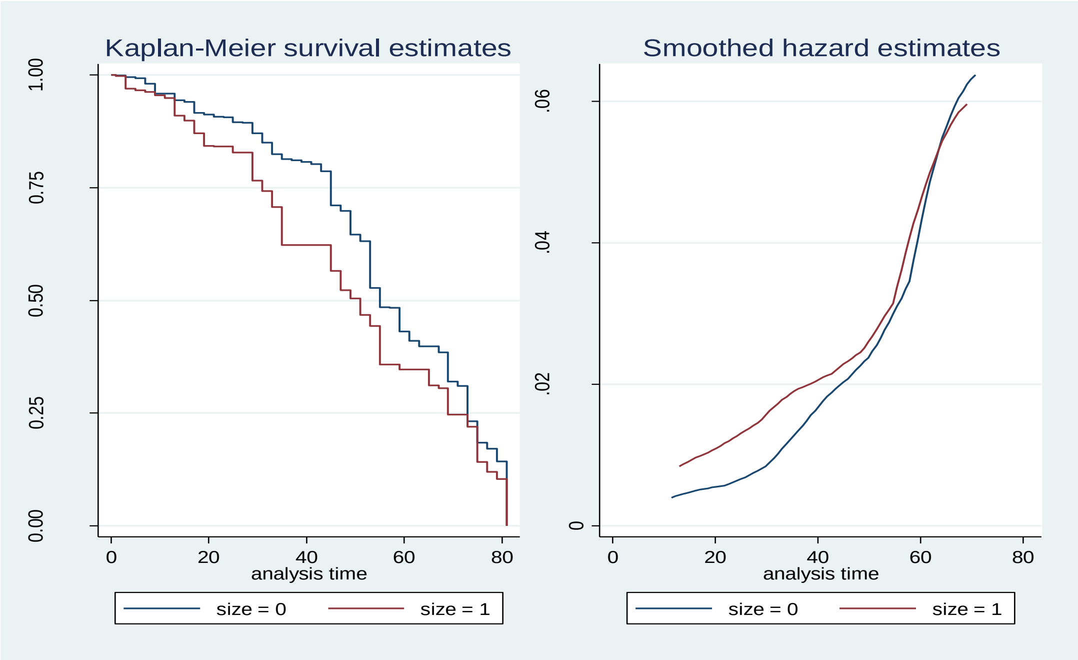
[8] However, using zero as a cut-off point for investment rates does not alter our results.

[9] Hypothesis tests based on the Log-Rank (or Generalized Savage), the Generalized Wilcoxon-Breslow and the Tarone-Ware confirm the results.

[10] Using the exponentially extended function does not alter the survival patterns in the zone of inactivity.

at the jth spell. For completeness and more clarity, our investment spike is defined as investment rates in excess of 20 percent. The model is estimated for investment in PME and plotted below by establishment size (size < 50 workers or size = 0) and large (size ≥ 50 workers or size =1).

Figure 3: Kaplan-Meier Survival and Hazard Estimates of Investment



The right panel in Figure 4.3 plots the empirical hazard expressed in Eq. 4 against the time since the last investment. It shows that the probability of having a lumpy investment episode is scale-dependent, where larger firms have a relatively higher probability of an investment spike compared to smaller firms. The hazard is increasing in the time since the last investment spike. Specifically, the hazard distribution is relatively flat initially and its slope becomes steeper soon thereafter reflecting increasing expenditure in M&R. Note that its shape is independent of whether the threshold investment spike used is 20 percent or 10 percent. The most striking result though is that the highest probability of a PME investment spike is less than 0.07 in the time elapsed since the last spike episode in Swaziland.[11]

This pattern of investment is consistent with an initially timid manufacturing sector seeking to wait until the uncertainty brought about by trade liberalization and entry/exit dynamics settles. In the

[11] This may appear to compare unfavourably with the probability of an investment spike of 0.66 for the USA in the year immediately succeeding an investment spike, 0.40 for Norway, 0.55 for Mexico and 0.60 for Colombia, (see Cooper *et al.*, 1999; Nielson and Schiantarelli, 2003; and Gelos and Isgut, 2001). All these studies consider only investment in equipment. Otherwise, the probability of plant acquisitions is likely to be lower than the probability of machinery procurement, while equipment purchasing is likely to occur more frequently than either plant or machinery transactions due to varying degrees of irreversibility. Thus, the probability of a spiky investment in all three asset classes combined will be reduced by the infrequent occurrence of investment in new physical plant.

process, depreciation and obsolescence of capital assets prevailed while their M&R increasingly became necessary during this period. Thus, and consistent with Cooper *et al*. (1999), this pattern of the hazard was primarily driven by the dominance of within-firm rather than between-firm effects in PME investments.

4. Econometric Models and Estimators

This section presents models and associated estimators that are useful in determining the probability of investing in durable capital goods in an environment of high inactivity. It uses several estimators as an unavoidable choice in a quantitative study based on an unbalanced dataset that has never been used before. Such dataset may potentially have measurement issues that are not yet fully understood. Therefore, the comparison of estimates obtained from different approaches is designed to rigorously scrutinize the robustness of estimates produced by any one of the estimators. Consequently, the section distinguishes between methods based on continuous and discrete responses according to how they handle initial conditions and endogeneity problems. It also makes a distinction between longitudinal dependence caused by the effects of preceding responses on succeeding responses and dependence arising from unobserved heterogeneity. In each of the modelling approaches, any setbacks related to estimation and potential solutions are discussed.

In that regard, three methods are outlined and discussed. *First*, for continuous responses, the long tradition of GMM approaches in estimating dynamic panel data models dominates empirical research, see Arellano and Bond (1991) and Blundell and Bond (1998). Kiviet *et al*. (2017) provide an extensive yet accessible exposition of the accuracy and efficiency of various GMM techniques in these models. *Second*, in the case of binary response models, a distinction between true state dependence and unobserved heterogeneity is achieved through dynamic modelling that includes a lagged response and a random intercept. The multilevel framework of analysis can be used to investigate the problem of these responses by constructing a joint model of the initial response with subsequent responses (e.g. Heckman, 1981a) and a model that conditions on the initial response (e.g. Wooldridge, 2005). A nonparametric maximum likelihood estimation (NPMLE) method due to Heckman and Singer (1984) can also be used in categorical responses. In both continuous and binary models, the assumption is that agents do not sort according to whether an agent operates in a high or low participation regime. *Third*, the final model therefore closes this gap by distinguishing between agents in high and low regimes of participation, see Lee and Frost (1978), Maddala (1983) and Lokshin and Sajaia (2004).

4.1. The GMM Approach

The linear dynamic panel data (DPD) model to be estimated is of the form

$$y_{it} = \boldsymbol{\alpha} y_{it-1} + \boldsymbol{\beta x} + \eta_i + v_{it}, \tag{5}$$

for $i = 1, \dots, N$, and $t = 2, \dots, T$, $\boldsymbol{\beta}$ a vector coefficients and $\boldsymbol{x}$ a vector of covariates, where a large N and small T DPD structure are assumed. The measure of state dependence $|\boldsymbol{\alpha}| < 1$[12] ensures convergence of the system, where η_i denotes individual-specific effects and v_{it} is the random error term. Arellano and Bond (1991) start with a first-order autoregressive – AR(1) − version of Eq. 5 that excludes the vector of strictly exogenous variables, $\boldsymbol{x}_{it}$

$$y_{it} = \boldsymbol{\alpha} y_{it-1} + u_{it} \tag{6}$$

where $u_{it} = \eta_i + v_{it}$ is the standard one-way error component structure representing fixed effects and random noise. The expected values of η_i and v_{it} are assumed equal to zero and $E(\eta_i v_{it}) = 0$ for $i = 1, \dots, N$ and $t = 2, \dots, T$. It is also assumed that $E(v_{it} v_{is}) = 0$ for $t \neq s$ and initial conditions satisfy $(y_{i1} v_{it}) = 0$. Taking first-differences (FD) of Eq. 6 yields

$$\Delta y_{it} = \ \boldsymbol{\alpha} \Delta y_{it-1} + \Delta u_{it}. \tag{7}$$

The $\boldsymbol{\alpha}$-coefficient of the lagged response, y_{it-1}, is the parameter of interest and measures the influence of the lagged response on the current behaviour of the dependent variable.

4.1.1 The Difference−GMM

The moment restrictions above are associated with $\frac{1}{2}(T-2)(T-1)$ linear orthogonality conditions in parameters for the GMM estimator; see Arellano and Bond (1991), Blundell and Bond (1998), and Bun and Windmeijer (2010). Using the notation of Bun and Windmeijer (2010), and Hayakawa and Pesaran (2015), it is assumed that

$E(y_i^{t-2} \Delta u_{it}) = 0$ for $t = 3, \dots, T$, where $y_i^{t-2} = (y_{i1}, y_{i2}, \dots, y_{it-2})'$ and $\Delta u_{it} = u_{it} - u_{it-1} = \Delta y_{it} - \ \boldsymbol{\alpha} \Delta y_{it-1}$.

The resultant sparse instrument matrix for the i^{th} firm, $\boldsymbol{Z}_{D,i}$, is then constructed as

$$\boldsymbol{Z}_{D,i} = \begin{pmatrix} y_{i1} & 0 & 0 & 0 & 0 & 0 & 0 \ \dots \ 0 \ \dots \ 0 \\ 0 & y_{i1} & y_{i2} & 0 & 0 & 0 & 0 \ \dots \ 0 \ \dots \ 0 \\ 0 & 0 & 0 & y_{i1} & y_{i2} & y_{i3} & 0 \ \dots \ 0 \ \dots \ 0 \\ . & . & . & . & . & . & 0 \ \dots \ 0 \ \dots \ 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \ \dots \ y_{i1} \cdots y_{iT-2} \end{pmatrix}$$

[12] See Hayakawa (2009, 2014) for a large N and large T DPD model.

where the set of linear moment conditions gives rise to an asymptotically efficient GMM that minimizes the following GMM criterion function, which is in turn based on Hansen (1982):

$$J(\hat{\alpha}) = \left(\sum_{i=1}^{N} \Delta u_i' Z_{D,i}\right) W_N \left(\sum_{i=1}^{N} Z_{D,i}' \Delta u_i\right).$$

The associated GMM estimator for α is given by Arellano and Bond (1991) and presented here as

$$\hat{\alpha}_{Diff} = \frac{\Delta y_{-1}' \boldsymbol{Z}_{D,i} W_N' \boldsymbol{Z}_{D,i}' \Delta y}{\Delta y_{-1}' \boldsymbol{Z}_{D,i} W_N' \boldsymbol{Z}_{D,i}' \Delta y_{-1}}$$

where $\Delta y = (\Delta y_1', \Delta y_2', \dots, \Delta y_N')'$, $\Delta y_i = \Delta y_{i3,} \Delta y_{i4,} \dots, \Delta y_{iT}$ and $\boldsymbol{Z}_d = (Z_{d1}', Z_{d2}', \dots, Z_{dN}')'$ and W_N is a two-step weighting matrix assuring validity of efficiency properties for the GMM estimator. The matrix is defined as

$$W_N = W_{Two-Step} = \left(\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{Z}_{D,i}' \Delta\hat{u}_i \, \Delta\hat{u}_i' \boldsymbol{Z}_{D,i}\right)^{-1}$$

Similarly, the one-step weighting matrix is given by

$$W_{One-Step} = \left[\frac{1}{N}\sum_{i=1}^{N} (Z_i' H Z_i)\right]^{-1}$$

and does not depend on estimated parameters. The square matrix H is of a $(T-2)(T-2)$ dimension with 2s on the main diagonal, -1s on the immediate off-diagonal and zeroes elsewhere (see Bond, 2002).

A few observations concerning the GMM DIFF estimator, $\hat{\alpha}_{Diff}$, need to be made. *First*, notice that W_N depends on parameter estimates through $\Delta\hat{u}_i = \Delta y_{it} - \hat{\boldsymbol{\alpha}}\Delta y_{it-1}$ and causes a downward bias on the estimated asymptotic standard errors of the two-step $\hat{\alpha}_{Diff}$, see Alonso-Borrego and Arellano (1999), Ziliak (1997) and Altonji and Segal (1996). Using the Taylor series expansion, Windmeijer (2005) identifies the source of bias and provides corrected asymptotic standard errors for the two-step GMM estimator. *Second*, as $T \to \infty$, the number of orthogonality conditions increases. Since growth in the number of moment conditions is quadratic in T, this leads to an explosion of instrument count. The standard solution to this in empirical studies involves collapsing the instrument set and/or curtailing its lag depth. *Thirdly*, in applications with persistent series where $\hat{\alpha}$ is near unity, for which the System GMM is more suitable, the process takes long to decay (see Roodman, 2009b and Han and Philips, 2010). It might also be the case that $(\sigma_{\eta i}/\sigma_{vi}) \to \infty$, implying a random walk with firm-specific drifts, creating weak correlations between first differences and lagged levels, or the weak instruments problem (see Blundell and Bond, 2000:325).

4.1.2 The System GMM

The unsatisfactory performance of the two-step differenced GMM estimator prompted Blundell and Bond (1998) to develop an estimator initially proposed by Arellano and Bover (1995). These authors proposed a System GMM estimator in which the moment conditions allow for the joint use of DIFF and LEV to circumvent the weak instruments problem and enhance the efficiency of the estimator. This required restrictions on the initial conditions and the assumption that

$$E(\eta_i \Delta y_{i2}) = 0$$

which holds when the process is mean-stationary (see Bun and Windmeijer, 2010) as

$$y_{i1} = \frac{\eta_i}{1-\alpha} + \varepsilon_i$$

where $E(\varepsilon_i) = E(\eta_i \varepsilon_i) = 0$. If the regularity conditions above hold, then $\frac{1}{2}(T-1)(T-2)$ moment conditions below are valid

$$E\left(u_{it} \Delta y_i^{t-1}\right) = 0$$

where $\Delta y_i^{t-1} = (\Delta y_{i2}, \Delta y_{i3}, \dots, \Delta y_{it-1})'$. With these moment conditions, it is possible to define a level's instrumental matrix as

$$\boldsymbol{Z}_{L,i} = \begin{pmatrix} \Delta y_{i2} & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \Delta y_{i3} & \Delta y_{i2} & 0 & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ . & . & . & . & . & . & . & \dots & . & \dots & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \Delta y_{iT-1} & \dots & \Delta y_{i2} \end{pmatrix}$$

together with $u_i = \begin{pmatrix} u_{i3} \\ u_{i4} \\ . \\ . \\ . \\ u_{iT} \end{pmatrix}$.

Following Bun and Windmeijer (2010), it is also true that

$$E\left(u_{it} \Delta y_i^{t-1}\right) = E(Z'_{Li} u_i) = 0.$$

Therefore, the levels−GMM estimator constructed from these moment conditions and $\boldsymbol{Z}_L$ is

$$\hat{\boldsymbol{\alpha}}_{LEV} = \frac{y_{-1}'\boldsymbol{Z}_L W_N^{-1} Z_L' y}{y_{-1}'\boldsymbol{Z}_L W_N^{-1}\boldsymbol{Z}_L' y_{-1}}.$$

Finally, the full set of moment conditions as supplied by Bun and Windmeijer (2010) based on the assumptions above can summarized as

$$\begin{cases} E\left(y_i^{t-2}\Delta u_{it}\right) = 0 \\ E\left(u_{it}\Delta y_i^{t-1}\right) = 0 \end{cases}$$

or

$$E(Z_{si}' p_i) = 0$$

and the instrumental matrix for calculating the system GMM is then given by

$$\boldsymbol{Z}_{SYS,i} = \begin{pmatrix} Z_{D,i} & 0 & 0 & \dots & 0 \\ 0 & y_{i2} & 0 & \dots & 0 \\ 0 & 0 & y_{i3} & \dots & 0 \\ . & . & . & \dots & 0 \\ 0 & 0 & 0 & \dots & y_{iT-2} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Z}_{D,i} & 0 \\ 0 & \boldsymbol{Z}_{L,i} \end{pmatrix}$$

$$\text{and } \; p_i = \begin{pmatrix} \Delta u_i \\ u_i \end{pmatrix}.$$

The systems−GMM estimator based on the full set of moment conditions is given by

$$\hat{\alpha}_{SYS} = \frac{q_{-1}'\boldsymbol{Z}_s W_{NS}^{-1} Z_s' q}{q_{-1}'\boldsymbol{Z}_s W_{NS}^{-1}\boldsymbol{Z}_s' q_{-1}}$$

where $q_i = (\Delta y_i', y_i')'$. In this case, the weighting matrix is given by

$$W_{NS} = \left(\frac{1}{N}\sum_{i=1}^{N} Z_{SYS}' M Z_{SYS}\right)^{-1}$$

where $M = \begin{bmatrix} H & 0 \\ 0 & I_{T-1} \end{bmatrix}$ or, as in Blundell and Bond (1998), $M = \begin{bmatrix} I_{T-1} & 0 \\ 0 & I_{T-1} \end{bmatrix} = I_{2T-2}$ with I_{T-1} representing an identity matrix.

When these conditions are met, the system GMM estimator has better finite sample properties than the differenced GMM estimator in terms of bias and root mean squared error (RMSE), see Blundell and Bond (1998) and Blundell, Bond and Windmeijer (2000).[13]

4.1.3 Forward Orthogonality Deviations, First Differences Transform and Instrument Proliferation

The first-difference transform has a specific weakness in that data gaps are magnified, especially in unbalanced panels. For example, suppose y_{it} is missing, then Δy_{it} and Δy_{it+1} are missing as well. This problem was first motivated by Arellano and Bover (1995) who developed a forward orthogonal deviations' operator that subtracts the average of all future values of the variable of interest. As an alternative to the FD routine, the orthogonal deviations transform is usefully applicable in models with predetermined regressors. The construction of the transform is explained in Arellano and Bover (1995:41) and simplified in Roodman (2009a). It relies on the Helmert's transformation for the variable ω formulated as

$$\omega_{i,t+1}^{\perp} = c_{it}\left(\omega_{it} - \frac{1}{T_{it}}\sum_{s>t}\omega_{is}\right)$$

where the scale factor, c_{it}, is chosen such that $c_{it} = \sqrt{\frac{T_{it}}{(T_{it}+1)}}$.[14] The term in brackets measures the deviations of each ω_{it} from the mean of its $T-1$ remaining future values. For an unbalanced dataset, the forward deviations operator is

$$A = diag\left[\frac{T-1}{T}, \ldots, \frac{1}{2}\right]^{1/2} x$$

$$\begin{bmatrix} 1 & -(T-1)^{-1} & -(T-1)^{-1} & \ldots & -(T-1)^{-1} & -(T-1)^{-1} & -(T-1)^{-1} \\ 0 & -1 & -(T-2)^{-1} & \ldots & -(T-2)^{-1} & -(T-2)^{-1} & -(T-2)^{-1} \\ . & . & . & \ldots & . & . & . \\ . & . & . & \ldots & . & . & . \\ . & . & . & \ldots & . & . & . \\ 0 & 0 & 0 & \ldots & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -1 \end{bmatrix}.$$

[13] The problem of high autoregressive parameter; $\hat{\boldsymbol{\alpha}} \rightarrow 1$ and $(\sigma_{\eta i}/\sigma_{vi}) \rightarrow \infty$, leading to the weak instruments problem also characterizes the SYS GMM estimator (*see* Bun and Windmeijer, 2010 and Han and Philips, 2010). Econometric theorists making propositions for optimizing the parametric efficiency of the SYS GMM include Bun and Windmeijer (2010), Han and Philips (2010), Youssef *et al*. (2014), Youssef and Abonazel (2015) and Kiviet *et al*. (2017).

[14] Demeaning the data prior to the Helmert transformation has no effect on the final results, see Appendix A4.2 for details.

In the case of a balanced dataset, for example, Roodman (2009a) provides an operator for the forward orthogonal transform typically expressed as $I_N \otimes M_{\perp}$, where

$$M_{\perp} = \begin{Bmatrix} \sqrt{\frac{(T-1)}{T}} & \frac{1}{\sqrt{T(T-1)}} & \frac{1}{\sqrt{T(T-1)}} & \cdots \\ & \sqrt{\frac{(T-2)}{(T-1)}} & \frac{1}{\sqrt{(T-1)(T-2)}} & \cdots \\ & & \sqrt{\frac{(T-3)}{(T-2)}} & \cdots \end{Bmatrix}$$

In this transformation, the rows of $M_{\perp}$ are orthogonal to each other. This means that ω_{it} remains independently distributed even after the transformation. The choice of c_{it} ensures that ω_{it} is also i.i.d.; i.e. $M_{\perp}M_{\perp}' = \mathrm{I}$. This is an expression portraying the assumption of homoscedasticity carried out in Arellano and Bond (1991).

It remains a concern that the GMM approach suffers from instrument proliferation arising from the increase in moment conditions as T increases, see Tauchen (1986), Ziliak (1997), Altonji and Segal (1996) and Bowsher (2002). In the discussion by Roodman (2009b) and Kiviet *et al*. (2017), the excessive number of instruments over-fit endogenous variables, produce imprecise estimates of the optimal weighting matrix, bias the two-step standard errors downward and weaken the Hansen Test of instrument validity. When instrument explosion characterizes the analysis, there are three standard methods for reducing the instrument count: (1) truncation of the lag depth of endogenous explanatory variables, (2) collapsing the instrument matrix (see Roodman, 2009a) and (3) both truncation of lag length and collapsing of instrument matrix. A new technique based on the principal component analysis has been theoretically analysed by Kapetanios and Marcellino (2007), Bai and Ng (2010) and Mehrhoff (2009) and has been empirically developed by Bontempi and Mammi (2015).

4.2 Nonlinear Dynamic Random-Effects Models and Estimators

4.2.1 The Multilevel Model

The GMM approach relies on continuous responses when treating state dependence and initial conditions in DPD models. In order to distinguish between the effects of true state dependence and unobserved heterogeneity on investment rates, we use a dichotomous dynamic response model that incorporates a lagged response and a firm-specific random-intercept. Three approaches to treating the initial conditions problem are adopted: (1) joint modelling of initial and subsequent responses using the one-factor model of Aitkin and Alfo (2003), (2) conditional modelling of subsequent responses given initial conditions and (3) the nonparametric maximum likelihood estimation (NPMLE).

Specifically, we draw heavily from Skrondal and Rabe-Hesketh (2014) who provide extensions of joint and conditional approaches. This method presents the probability of an outcome of the response variable using the standard assumption of normally distributed idiosyncratic shocks and the random-intercept term as

$$\begin{cases} Pr\left(y_{ij} = 1 | y_{i-1,j}, \boldsymbol{z}_j, \boldsymbol{x}_{ij}, \zeta_j\right) = h^{-1}\left(\boldsymbol{z}_j' \boldsymbol{\gamma_z} + \boldsymbol{x}_{ij}' \boldsymbol{\gamma_x} + \boldsymbol{\alpha} y_{i-1,j} + \zeta_j\right) \\ Pr\left(y_{0j} = 1 | \boldsymbol{z}_j, \boldsymbol{x}_j, \zeta_j\right) = h^{-1}\left(\boldsymbol{z}_j' \boldsymbol{g_z} + \boldsymbol{x}_{0j}' \boldsymbol{g_x} + \lambda_0 \zeta_j\right) \end{cases}, \quad i = 1, 2, \dots, T-1 \tag{8}$$

$$\boldsymbol{y_{ij}^*} = \boldsymbol{z}_j' \boldsymbol{\gamma_z} + \boldsymbol{x}_{ij}' \boldsymbol{\gamma_x} + \boldsymbol{\alpha} y_{i-1,j} + \zeta_j + \varepsilon_{ij} \tag{8)$'$(}$$

where $\zeta_j \sim \mathbb{N}(0, \psi)$, j=1, …, N, and $\boldsymbol{\gamma_z}$ and $\boldsymbol{\gamma_x}$ are the coefficient vectors for the time-invariant $\boldsymbol{z}_j$ and time-varying $\boldsymbol{x}_{ij}$ covariates, respectively. The link function h(·) is a probit function linking the conditional expectation of y_{ij} to the linear predictor on the right-hand side; that is, $h\left(p_{ij}\right) = \Phi^{-1}\left(p_{ij}\right)$, where $p_{ij} = \mathrm{pr}\left(y_{ij} = 1 | y_{i-1,j}, \boldsymbol{z}_j, \boldsymbol{x}_{ij}, \zeta_j\right)$. Eq. 8 can be expressed in latent form as in Eq. $(8)'$. Here the threshold model connects observed responses to latent responses as $y_{it} = I(\mathbf{y_{it}^*} > 0)$ and $y_{i1} = I(\mathbf{y_{i1}^*} > 0)$. The indicator function, $I(\cdot)$, takes the value of 1 if the expression in the bracket holds and 0 otherwise. In this case, the firm-specific random-effects specification used here implies that the correlation between the total error component, $u_{it} = \zeta_j + \varepsilon_{ij}$, in any two different occasions is constant: $\frac{\psi}{\psi+1}$.

In Skrondal and Rabe-Hesketh (2014), a one-factor component with occasion-specific factor loading λ_i is introduced to the right-hand side of Eq. 8. This factor model for binary responses is naturally restricted to have one free factor loading λ_0 for the initial response and $\lambda_i = 1$ for the subsequent responses. In order to control for level 2 endogeneity of x_j in the one-factor model, we follow the standard practice due to Mundlak (1978) and Chamberlain (1984) by using the auxiliary model

$$\zeta_j = \delta_{\bar{x}_{.j}} \bar{x}_{.j} + u_j$$

where $u_j \sim N(0,1)$ is independent of $\bar{x}_{.j}$. Chamberlain (1984) observes that in nonlinear random-intercept models, the auxiliary equation represents a proper statistical model which must be correctly specified. Thus the use of $\bar{x}_{.j}$ instead of x_j restricts the correlations between the random-intercept and the time-varying covariates to be constant over time.

When x_{ij} has missing values, using longitudinal means is usually the only viable option in practice, see Rabe-Hesketh and Skrondal (2012). In that case, the calculation of $\bar{x}_{.j}$ is based on only those occasions for which the response variable y_{ij} contributes to the analysis. Substituting the auxiliary equation in Eq. 8, the linear model of latent responses becomes

$$\begin{cases} \Pr\left(y_{ij} = 1|y_{i-1,j}, \boldsymbol{z}_j, \boldsymbol{x}_{ij}, \zeta_j\right) = h^{-1}\left(\boldsymbol{z}_j'\boldsymbol{\gamma}_{\boldsymbol{z}} + \boldsymbol{x}_{ij}'\boldsymbol{\gamma}_{\boldsymbol{x}} + \overline{\boldsymbol{x}}'_{.j}\boldsymbol{\delta}_{\overline{\boldsymbol{x}}} + \boldsymbol{\alpha}y_{i-1,j} + \mathrm{u}_j\right) \\ \quad \Pr\left(y_{0j} = 1|\boldsymbol{z}_j, \boldsymbol{x}_j, \zeta_j\right) = h^{-1}\left(\boldsymbol{z}_j'\boldsymbol{g}_{\boldsymbol{z}} + \boldsymbol{x}_{0j}'\boldsymbol{g}_{\boldsymbol{x}} + \overline{\boldsymbol{x}}'_{.j}\lambda_0\boldsymbol{\delta}_{\overline{\boldsymbol{x}}} + \lambda_0\mathrm{u}_j\right) \end{cases}, \text{ i=1, …, T-1} \quad (9)$$

Again, in order to handle level 2 endogeneity, the conditional modelling approach used is

$$\zeta_j = \delta_y y_{0j} + \boldsymbol{z}_j'\delta_z + \boldsymbol{x}_{0j}'\delta_{x0} + \overline{\boldsymbol{x}}'_{.j}\boldsymbol{\delta}_{\overline{\boldsymbol{x}}} + u_j \quad (10)$$

where the longitudinal averages can be calculated according to Rabe-Hesketh and Skrondal (2013) as

$$\bar{x}_i = \frac{1}{T-1}\sum_{1}^{T} x_{it}$$

and a probit link in Eq. 9 is maintained.

4.2.2 The Nonparametric Maximum Likelihood Estimator

In Heckman and Singer (1984), a nonparametric maximum likelihood estimation (NPMLE) procedure that avoids *ad hoc* functional specifications for the unobserved scalar heterogeneity θ is proposed. The nonparametric characterization of the marginal density of investment $f(y_i|X_i)$ becomes

$$f(y_i|X_i) = \sum_{j=1}^{k} g\left(y_i|X_i, \theta_j\right)p_j$$

where $\sum p_j = 1, p_j \geq 0, j = 1, \dots, k, k$ is the number of points of support, p_j is probability mass point, θ_j is a locator of p_j such that $p_j = \mathrm{prob}\left(\theta = \theta_j\right)$. Under random sampling, the log-likelihood for investment rates is given by

$$LL = \sum_{i=1}^{N} ln \sum_{j=1}^{k} g(y_i|X_i, \theta_i)\, p_j.$$

Lindsay (1983) provides conditions for global solution to the maximization of LL using the Gateaux variation. The Gateaux derivative of the log-likelihood function with respect to θ is defined as

$$D(\theta, \mu) = \sum_{i=1}^{M}\left[\frac{g\left(y_i|X, \theta_j\right)}{f\left(y_i|X_i\right)} - 1\right].$$

The log-likelihood function is maximized if and only if $D(\theta, \mu) \leq 0$ for all $\theta_j \in \Theta$, see the Mass Point Method section in Huh and Sickles (1994). Heckman and Singer (1984) derive $\theta \in [\theta_{min}, \theta_{max}]$ over which $g\left(y_i|X_i, \theta_i\right)$ is supported. The Heckman-Singer estimator has been found consistent for mixing distributions with a small number of points of support.

4.3 Endogenous Switching Regression Model of Investment

DPD models of investment estimated using the GMM approach or multilevel methods assume that an optimal rate of investment is characterized by a single investment regime. For example, Abel and Eberly (1994) and Abel (2014) demonstrate that the optimal rate of investment can be located in more than one regime. In such environments, micro investment decisions concern not only whether a firm invests, but also how much it invests in a regime. The goal here is to estimate the switching regression model specified in Eq. 11

$$\begin{cases} \frac{I_{it}}{K_{it-1}} = X_{it}\beta^{Low} + \varepsilon_{1it} & \text{iff} \quad Z_{it}\gamma + u_{it} < 0 \\ \frac{I_{it}}{K_{it-1}} = X_{it}\beta^{High} + \varepsilon_{2it} & \text{iff} \quad Z_{it}\gamma + u_{it} \geq 0 \end{cases} \tag{11}$$

and

$$\begin{pmatrix} u_1 \\ u_2 \\ \varepsilon \end{pmatrix} \sim \mathbb{IN}(\mathbf{0}, \mathbf{\Sigma}), \quad \text{with} \quad \mathbf{\Sigma} = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{1\varepsilon} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{2\varepsilon} \\ \sigma_{\varepsilon 1} & \sigma_{\varepsilon 2} & 1 \end{pmatrix}$$

where the Z_{it} vector includes variables in the switching regression function and an additional variable to operate as an exclusion restriction to correct for selection bias, see Cameron and Trivedi (2009).

The non-zero covariance between investment shocks $\varepsilon_{1it}, \varepsilon_{2it}$ and u_{it} in Eq. 11 is correlated with other firms' characteristics. Since the conditions that either $\varepsilon_{1it} \neq 0$ or $\varepsilon_{2it} \neq 0$ or both are assumed to hold, then Eq. 11 is an endogenous switching regression model. Investment rates observed in each period *t* for each firm *i* are generated from either the High-q or Low-q regime, but never in both at any one time. As a consequence, the covariance between ε_{1it} and ε_{2it} does not exist, see Maddala (1983). By definition, the vector $X_{it} = f(q_{it})$ is a set of observable exogenous explanatory variables. As in Lee and Porter (1984) and Hu and Schiantarelli (1998), it is unknown *ex ante* whether the observed investment rate is generated from the High-q or Low-q regime. That is, unlike Nabi (1989), we have a case of unknown sample separation in the model.

5 Empirical Results

We now take our GMM estimators, dynamic nonlinear random effects models and endogenous regime switching models to the Swazi manufacturing panel data. As argued earlier, our preferred sufficient statistic that measures or poses as a proxy for marginal q is the sales-to-capital ratio. Another covariate is the time $t-1$ investment rate accommodating the conditional probability of a positive investment in the future as a function of previous investment that captures investment dynamics; see Heckman (1981b). It is also standard in state dependence research to control for unobserved heterogeneity. We therefore control for individual characteristics underlying the firm's decision to either invest or exercise its option to wait. In view of the argument presented by Hsiao (2003) and Chrysanthou (2008) that state dependence and unobserved heterogeneity have opposite effects on firms' investment decisions, it is necessary to determine the relative importance of each one of them. Finally, the empirical estimation strategy takes into account the likelihood of capital/labour substitutability in production by introducing employment as a control variable.

The linear DPD in Eq. 5 can therefore be specified as a structural empirical model of investment in the form shown in Eq. 12

$$\frac{\boldsymbol{Investment_{it}}}{\boldsymbol{Capital_{it-1}}} = \boldsymbol{\alpha}\left(\frac{\boldsymbol{Investment_{it-1}}}{\boldsymbol{Capital_{it-2}}}\right) + \boldsymbol{\beta_1}\left(\frac{\boldsymbol{Sales_{it}}}{\boldsymbol{Capital_{it-1}}}\right) + \boldsymbol{\beta_2}\left(\frac{\boldsymbol{Sales_{it-1}}}{\boldsymbol{Capital_{it-2}}}\right) + \boldsymbol{\beta_3}(\mathrm{Emp}_{\boldsymbol{it}}) + \boldsymbol{\beta_4}(\mathrm{Emp}_{\boldsymbol{it-1}}) + \boldsymbol{u_{it}} \quad (12)$$

where the two-way error structure is defined as $\boldsymbol{u_{it}} = \boldsymbol{\mu_i} + \boldsymbol{\tau_t} + \boldsymbol{\varepsilon_{it}}$, for $t = 2, \ldots, T$, $\boldsymbol{\mu_i}$ and $\boldsymbol{\tau_t}$ are the unobservable firm-specific effect and time effects, respectively; while $\boldsymbol{\varepsilon_{it}}$ is the random error term. The dependant variable is the rate of investment in PME in the manufacturing sector. Its lagged regressor measures the state dependence of investment on the producer's previous decisions to invest. The contemporaneous sales-to-capital ratio included as a proxy for marginal q, while employment controls for primary input substitution effects.

The empirical DPD literature is awash with evidence of upward-biasedness of the OLS estimator when applied to Eq.12; that is, $\boldsymbol{\alpha} < \mathrm{plim}\ (\widehat{\boldsymbol{\alpha}}_{OLS}) < 1$, see Blundell, Bond and Windmeijer (2000). This inconsistency arises from, *inter alia*, the correlation of the lagged level of the dependent variable, $\frac{\boldsymbol{Investment_{it-1}}}{\boldsymbol{Capital_{it-2}}}$, with the stochastic error term $\boldsymbol{\varepsilon_{it}}$, see Bond (2002). According to Judson and Owen (1999), the bias of $\widehat{\boldsymbol{\alpha}}_{OLS}$ is much more severe than that of $\widehat{\boldsymbol{\beta}}$. Similarly**,** the within-groups (WG) estimator has proved to be downward-biased. Alvarez and Bond (2003) demonstrate that the asymptotic bias in GMM is always smaller than the bias in WG, provided $T < N$. Therefore, a consistent estimator $\widehat{\boldsymbol{\alpha}}$ should be bounded below and above such that $\widehat{\boldsymbol{\alpha}} \in (\widehat{\boldsymbol{\alpha}}_{WG}, \widehat{\boldsymbol{\alpha}}_{OLS})$. The rest of the results are estimated using the GMM technique, methods for the estimation of dynamic nonlinear random effects models with unbalanced panel data and the endogenous investment switching regime approach.

5.1 The GMM Estimates

5.1.1 The *Difference* and *System* GMM Results

Judson and Owen (1999) propose that when $T = 10$ and $N > 100$, Difference and System GMM should be used in estimating DPD models. Moreover, the added advantage of the System GMM is that it performs better than the Difference GMM in applications with near unit-root time series data. In such cases, lagged levels of variables are weak instruments for subsequent variations – see Roodman (2009b), Blundell and Bond (1998, 2000), and Blundell *et al*. (2000).

Table 3 summarizes the empirical anatomy of section 4. The first column characterizes the GMM parameters, $\beta \in [\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4]$. These are estimated using the One-Step and Two-Step approaches of the Difference and System GMM. The parameter estimate, $\hat{\alpha}$, denotes the estimated lagged response coefficient and the rest are coefficients of other explanatory variables that may be assumed endogenous, predetermined or strictly exogenous.[15] However, moment conditions by Arellano and Bond (1991) for Difference GMM and by Blundell and Bond (1998) for System GMM ensure asymptotic consistency of parameters.

Table 3: Schema for the GMM Estimator Using Arellano and Bond (1991) for $\hat{\alpha}_{Diff}$ and Blundell and Bond (1998) for $\hat{\alpha}_{SYS}$

Parameter	$\hat{\boldsymbol{\alpha}}_{\boldsymbol{Diff}} = \frac{\Delta y'_{-1}\boldsymbol{Z}_{D,i}W'_N\boldsymbol{Z}'_{D,i}\Delta y}{\Delta y'_{-1}\boldsymbol{Z}_{D,i}W'_N\boldsymbol{Z}'_{D,i}\Delta y_{-1}}$		$\hat{\boldsymbol{\alpha}}_{\boldsymbol{SYS}} = \frac{q'_{-1}\boldsymbol{Z}_s W_{NS}^{-1} Z'_s q}{q'_{-1}\boldsymbol{Z}_s W_{NS}^{-1}\boldsymbol{Z}'_s q_{-1}}$	
	One−Step	Two−Step	One−Step	Two−Step
$\hat{\alpha}$	$\hat{\alpha}_{1Diff}$	$\hat{\alpha}_{2Diff}$	$\hat{\alpha}_{1SYS}$	$\hat{\alpha}_{2SYS}$
$\hat{\beta}_1$	$\hat{\beta}_{1.1Diff}$	$\hat{\beta}_{1.2Diff}$	$\hat{\beta}_{1.1SYS}$	$\hat{\beta}_{1.2SYS}$
$\hat{\beta}_2$	$\hat{\beta}_{2.1Diff}$	$\hat{\beta}_{2.2Diff}$	$\hat{\beta}_{2.1SYS}$	$\hat{\beta}_{2.2SYS}$
$\hat{\beta}_3$	$\hat{\beta}_{3.1Diff}$	$\hat{\beta}_{3.2Diff}$	$\hat{\beta}_{3.1SYS}$	$\hat{\beta}_{3.2SYS}$
$\hat{\beta}_4$	$\hat{\beta}_{4.1Diff}$	$\hat{\beta}_{4.2Diff}$	$\hat{\beta}_{4.1SYS}$	$\hat{\beta}_{4.2SYS}$
Constant	−	−	$\hat{\beta}_0$	$\hat{\beta}_0$

The schema in Table 3 treats the model as a system of equations, one for each time period, as in Bontempi and Golinelli (2014). First, the predetermined and endogenous variables in first-differences are instrumented with suitable lags of their own levels. Second, predetermined and endogenous variables in levels are instrumented with suitable lags of their own first-differences. Lastly, strictly exogenous and any other instruments enter the instrument matrix with one column per instrument.

The empirical model in Eq.12 is first estimated by WG and OLS methods to get the estimate $\hat{\boldsymbol{\alpha}} \in (0.005,\ 0.531)$ of the true value for $\boldsymbol{\alpha}$.[16] Table 4 estimates Eq.12 to produce baseline results based on *a priori* considerations that investment is a function of previous period's investment decisions and marginal q; that is, it is state dependent. Theory argues that although marginal q is a sufficient statistic for investment rate, the sales/capital ratio is also a sufficient statistic for investment rate as discussed, see Caballero and Leahy (1996). This means that, since marginal q is unobservable, the sales/capital variable can be used as a regressor instead. Therefore the empirical equation expresses the investment rate as a function of its $t-1$ lag, the contemporaneous sales/capital ratio and its $t-1$ lag.

[15] See definitions in Appendix A1.
[16] Full results are available from the author on request.

Table 4 reports estimates of true state dependence of real investment rates $\left(\frac{I_{t-1}}{k_{t-2}}\right)$ as well as t and $t-1$ sales/capital ratio. To achieve this, the one-step and two-step GMM parameters for $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{Diff}}$ and $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{SYS}}$ are respectively presented. The first-order autoregressive parameter is high and falls above the upper limit provided by the OLS estimator; that is, $\widehat{\boldsymbol{\alpha}} \to 1$, [and it might also be the case that $(\sigma_{\eta i}/\sigma_{vi}) \to \infty$]. This implies a random walk with firm-specific drift, creating weak correlations between first differences and lagged levels, or the weak instruments problem (see Blundell and Bond, 2000:325, and Han and Phillips (2010)). It may be a reflection of an imprecisely measured parameter due to high correlation between the sales/capital variable and omitted variables and other factors. Such is a natural characteristic of the GMM DIFF estimator while the GMM SYS estimator circumvents the problem. The one-step GMM SYS estimator has the autoregressive parameter $\widehat{\boldsymbol{\alpha}} > 1$, rendering the system non-convergent. The two-step GMM estimator barely passes the AR (1) restriction; and Roodman (2009a) suggests that the validity of the model need not be readily accepted in such cases. Furthermore, standard errors are Windmeijer (2005) robust bias-corrected.

Table 4: GMM Estimation of Investment Rate Dynamics using an Instrument Reduction Technique and the Helmert's Transform[17]

	GMM DIFF (COLL)		GMM SYS (COLL)	
	One−Step	Two−Step	One−Step	Two−Step
$\frac{I_{t-1}}{k_{t-2}}$	0.872*	0.584	1.044**	0.855
	(0.4173)	(0.5609)	(0.4042)	(0.4367)
$\frac{s_t}{k_{t-1}}$	−0.774*	−0.607	−0.853*	−0.794*
	(0.3366)	(0.4079)	(0.3508)	(0.3917)
$\frac{s_{t-1}}{k_{t-2}}$	0.106	0.09	0.239	0.074
	(0.1572)	(0.2265)	(0.1717)	(0.1838)
Constant	−	−	0.702	0.853
	−	−	(0.3922)	(0.4386)
NT	103	103	172	172
N	44	44	69	69
AR(1)−*p-value*	0.035	0.205	0.037	0.082
AR(2)−*p-value*	0.105	0.227	0.12	0.128
Sargan −*p-value*	0.1306	0.1306	0.029	0.029
Hansen −*p-value*	0.1395	0.1395	0.233	0.233
#Z	18	18	21	21
#X	10	10	10	10
Wald χ^2 −Test	42.97	47.54	39.52	37.06
χ^2_p	0	0	0	0.0001
h	3	3	3	3

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.
Notes: All models include Year Dummies.

Moreover, these results are an outcome of instrument proliferation that is controlled for by first collapsing the instrument count and truncation of lag depth to $t-2$. Instrument explosion curtailed by both mechanisms reduces proliferation from 79 to 18 instruments for *Difference* GMM and from 100

[17] A robustness check based on Bontempi and Mammi's (2015) principal component analysis technique presents similar results in Appendix A4.1.

to 21 instruments for *System* GMM. Furthermore, the *a priori* estimates of the variance-covariance of the transformed errors given by the blocks of H were used alternately between h(2) and h(3). By design, this has no effect on the $\hat{\boldsymbol{\alpha}}_{\boldsymbol{Diff}}$ results, but h(3) has the effect of slightly increasing the size of $\hat{\boldsymbol{\alpha}}_{\boldsymbol{SYS}}$ as evident on the table. Among the existing methods for expunging fixed effects, the method of forward orthogonal deviations is preferred due to its resilience to the gaps' problem. Such problems might be exacerbated, for example, by the use of the standard first difference deviations transform, given the high incidence of missing values in the investment data.

It is therefore not possible to draw sound conclusions on whether or not there exists true state dependence in Swazi manufacturing investments based on the table 4 results. Since micro level investment is shown in the survival rate section to differ by firm size, estimating the same structural model by controlling for firm-level employment alters the results somewhat. This is shown in Table 5 where the empirical model is estimated in full with employment as an additional control variable for primary input substitutability.

Table 5: GMM Estimation of Investment Rate Dynamics with the Control Variable using an Instrument Reduction Technique and the Helmert's Transform

	GMM DIFF (COLL)		GMM SYS (COLL)	
Variables	One−Step	Two−Step	One−Step	Two−Step
$\frac{I_{t-1}}{k_{t-2}}$	0.329	0.198	0.338	0.144
	(0.3041)	(0.4454)	(0.3509)	(0.4684)
$\frac{s_t}{k_{t-1}}$	−0.279	−0.121	−0.26	−0.073
	(0.2722)	(0.4135)	(0.3378)	(0.5149)
$\frac{s_{t-1}}{k_{t-2}}$	0.086	0.099	0.068	0.118
	(0.1341)	(0.2337)	(0.1253)	(0.2168)
Emp_t	−0.02	0.083	−0.147	−0.147
	(0.2592)	(0.3723)	(0.158)	(0.2343)
Emp_{t-1}	0.349*	0.386	0.297	0.273
	(0.1696)	(0.2002)	(0.1728)	(0.2152)
Constant	−	−	−0.36	−0.475
	−	−	(0.4997)	(0.6846)
NT	103	103	171	171
N	44	44	68	68
AR(1)−*p-value*	*0.035*	*0.097*	*0.023*	*0.08*
AR(2)−*p-value*	*0.041*	*0.094*	*0.021*	*0.134*
Sargan −*p-value*	*0.1477*	*0.1477*	*0.2006*	*0.2006*
Hansen −*p-value*	*0.1939*	*0.1939*	*0.3242*	*0.3242*
#Z	25	25	29	29
#X	12	12	12	12
Wald χ^2 −Test	106.53	76.23	93.57	71.21
χ^2_p	0	0	0	0
h	3	3	3	3

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.
Notes: All models include Year Dummies.

Although the size of the AR(1) parameter is substantially reduced across all estimators, it remains insignificant within the (WG, OLS) bounds. However, only the two-step GMM SYS (COLL) estimator passes the AR(1) and AR(2) Arellano-Bond (1991) diagnostic tests while the rest do not. However, in spite of lag length truncation and collapsing of instruments, the Hansen tests of over-identifying restrictions and joint instrument validity remain unsatisfied. For instance, the Hansen *p-value* of 0.3242 far exceeds the cautionary *p-value* of 0.25 suggested by Roodman (2009a). It therefore may be the case that the moment conditions are still too high thereby rendering the Sargan/Hansen test weak.

One interpretation of these results partly suggests the absence of persistence in investments due; *inter alia*, to the over 70% incidence of investment inactivity during the period of trade reforms gleaned in Figure 4.2. The *dominant* zone of inactivity in the data is modelled to respond to the previous period's inactivity and the sale/capital variable as a theoretical sufficient statistic of investment. Remember, the correlation between $t-1$ investment rate and t sales/capital ratio in Table 2 is 0.71 while the correlation between the current investment rate and its lag is 0.61. Firstly, the introduction of the $t-1$ investment rate and the t sales/capital ratio as explanatory variables causes collinearity and imprecision in the parametric estimation of the structural model. Secondly, turning to the use of only the $t-1$ explanatory variables without controls worsens the precision of the estimates potentially due to the impact of serial correlation since the $t-1$ sales/capital ratio is correlated with the $t-2$ investment rate. This $t-2$ investment rate is in turn correlated with its subsequent level. Again, the coefficients are measured with significant imprecision.

Nonetheless, although insignificant, the measure of state dependence is consistent with the findings in the literature in terms of its sign and order of magnitude; see Eberly *et al*. (2012) and Drakos and Konstantinou (2013). Taken at face value, an increase in the ratio of investment/capital stock at $t-1$ in the two-step GMM SYS (COLL) estimator, *ceteris paribus*, is more likely to have a positive effect on the probability of investing at time t than otherwise. Both contemporaneous control covariates; that is, the proxy for marginal q and the employment variable, might have negative effects on current investment rates while the $t-1$ individual lags might positively affect the time t investment rate.[18] As discussed; however, this framework of analysis ignores the potential effect of serial correlation in the time-varying errors much against the objection advanced by Honoré and Kyriazidou (2000).

In general, one explanation of the apparent industrial lacklustre performance in Swaziland is that capital irreversibility due to market failures acted as an investment deterrent in the uncertain business environment during the two decades since the 1990s in the customs union.[19] Most firms in the active group chose to exercise their option to wait for uncertainty to come down while maintaining and repairing existing plant, machinery and equipment. Only a few of the active firms engaged in lumpy

[18] It is possible that the time t covariates are correlated with firm-specific effects in the one-way error structure, thereby generating simultaneity problems. However, their exclusion in favour of retaining the $t-1$ covariates does not alter our results.

[19] Market failures in this case may be driven by 'lemon effects' and capital specificity, see Abel, Dixit, Eberly and Pindyck (1996).

investments after spells on inactivity. In this sense, investment decisions could not *significantly* respond to changes in the ratio of sales/capital and to changes in employment. The next section performs robustness checks to the estimation of the theoretical model and the model with controls using a different deviations transform to the data set to answer this question.

5.1.2 Sensitivity Analysis of the GMM Results

This section estimates the empirical model allowing for the impact of gaps in the dataset created by missing investment values. In order to check the robustness of the results obtained using the forward deviations orthogonality transform without 'employment' in the previous section, table 6 implements the same model but this time uses the first-difference deviations transform. In the absence of the employment variable, the coefficient of the lagged investment rate variable increases as expected and is weakly significant in three out of four cases. The magnification effect of the first-difference deviations transform in the estimation of GMM parameters is evident and marginally raises the coefficients above those estimated with our preferred forward orthogonality transform. However, the AR(1) parameter still lies outside the $(\mathrm{WG}, \mathrm{OLS})$ bracket.

Table 6: GMM Estimation of Investment Dynamics using the Roodman (2009b) Method of Instrument Reduction with Standard First Difference Deviations Transform without Controls

	GMM DIFF		GMM SYS	
Variables	One−Step	Two−Step	One−Step	Two−Step
$\frac{I_{t-1}}{k_{t-2}}$	1.014*	0.692	1.237**	0.972**
	(0.4244)	(0.5766)	(0.4494)	(0.3308)
$\frac{s_t}{k_{t-1}}$	−0.955*	−0.74	−1.127*	−0.963**
	(0.4234)	(0.4959)	(0.4755)	(0.3739)
$\frac{s_{t-1}}{k_{t-2}}$	0.106	0.106	0.207	0.054
	(0.1749)	(0.2331)	(0.2033)	(0.1418)
Constant	−	−	1.055	1.080*
	−	−	(0.6158)	(0.4368)
NT	100	100	172	172
N	43	43	69	69
AR(1)−*p-value*	0.033	0.218	0.052	0.076
AR(2)−*p-value*	0.126	0.257	0.164	0.148
Sargan −*p-value*	0.172	0.172	0.1213	0.1213
Hansen −*p-value*	0.2414	0.2414	0.4336	0.4336
#Z	18	18	21	21
#X	10	10	10	10
Wald χ^2 −Test	33.61	31.25	30.6	28.37
χ^2_p	0.0002	0.0005	0.0007	0.0016
h	3	3	3	3

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.
Notes: All models include Year Dummies.

As a further robustness check, table 7 applies the first-difference transform to estimate the empirical structural equation with controls. Here we control for employment size but continue with the standard first-difference deviations transform to estimate the model. Although it is substantially reduced in absolute terms, the true state dependence coefficient is still insignificant at any level. Although the autoregressive coefficient remains statistically insignificant and positive, its increase might also increase the probability of investing at time t by the order of approximately 0.30. Both the sales-to-capital ratio and employment behave similarly to the preferred specification.[20] The orders of magnitude and signs of these results mimic the findings of Drakos and Constantinou (2013) for the Greek manufacturing sector, whose estimated state dependence according to the research by these authors is found to lie between 0.19 and 0.33 for a similar period of analysis.

Table 7: GMM Estimation of Investment Dynamics using the Roodman (2009b) Method of Instrument Reduction with Standard First Difference Deviations Transform with Controls

	GMM DIFF		GMM SYS	
Variables	One−Step	Two−Step	One−Step	Two−Step
$\frac{I_{t-1}}{k_{t-2}}$	0.482	0.336	0.455	0.301
	(0.2904)	(0.3532)	(0.3752)	(0.532)
$\frac{s_t}{k_{t-1}}$	−0.416	−0.193	−0.377	−0.26
	(0.2992)	(0.4574)	(0.3532)	(0.5876)
$\frac{s_{t-1}}{k_{t-2}}$	0.077	0.087	0.069	0.116
	(0.1353)	(0.2029)	(0.12)	(0.1923)
Emp_t	0.111	0.249	−0.136	−0.161
	(0.3511)	(0.3596)	(0.1994)	(0.2282)
Emp_{t-1}	0.523*	0.644*	0.271	0.252
	(0.2371)	(0.2932)	(0.2148)	(0.23)
Constant	−	−	−0.166	−0.124
	−	−	(0.5313)	(0.8102)
NT	100	100	171	171
N	44	44	68	68
AR(1)−*p-value*	*0.042*	*0.093*	*0.025*	*0.091*
AR(2)−*p-value*	*0.057*	*0.05*	*0.033*	*0.116*
Sargan −*p-value*	*0.1991*	*0.1991*	*0.1246*	*0.1246*
Hansen −*p-value*	*0.1722*	*0.1722*	*0.2605*	*0.2605*
#Z	25	25	29	29
#X	12	12	12	12
Wald χ^2 −Test	55.74	43.29	56.3	37.8
χ^2_p	*0*	*0*	*0*	*0.0002*
h	3	3	3	3

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.
Notes: All models include Year Dummies.

Overall, the GMM results predict that micro investment rates in Swazi manufacturing insignificantly influence their own levels positively in the next period, and remain insignificant even when employment is controlled for. This is robust to the choice of deviations transform used. The results further reveal that the impact of contemporaneous sale-to-capital ratio is negative and insignificant while its $t-1$ coefficient is positive although still insignificant. This same pattern of parametric behaviour obtains in the case of the control variable. Thus, investment performance is invariant to the

[20] These results are robust to using the average profit of capital defined by Abel and Blanchard (1986) as $\left(\frac{VA_{t-1}-W_{t-1}}{K_{t-2}}\right)$, to using cash-flow to capital ratio $\left(\frac{CF_{t-1}}{K_{t-2}}\right)$ and operating profit to capital ratio $\left(\frac{\pi_{t-1}}{K_{t-2}}\right)$ defined by Letterie and Pfann (2007). All three are considered as proxies for the shadow price of capital.

choice of a deviations' transform applied to the treatment of missing values. Are these conclusions sensitive to the treatment method applied to missing values of investment? Does an interaction between missingness patterns of values and employment variations has an effect on the rate of investment?

In this framework, the impact of firms' investment inactivity on industrial investment patterns can potentially be indirectly accounted for through variations in the orthogonality conditions assumed. The purging of individual fixed effects in the GMM approach removes information about plant-level heterogeneity in investment decisions. Browning and Carro (2010) and Skrondal and Rabe-Hesketh (2014) develop binary discrete choice models with heterogeneity as an important factor to take into account in inference analysis based on microdata. In the next section, we depart from modelling continuous responses of investment and introduce a binary method to estimating dynamic nonlinear random effects models of unbalanced panels of firms in a multilevel setting of investment.

5.2 Dynamic Random-Effects Estimates

5.2.1 Empirical Multilevel Analysis of Investment Decisions

The empirical version of the dynamic random-effects model follows directly from the theoretical specification and can be concisely summarized as

$$\text{probit}\left\{\left(\frac{\boldsymbol{I_{ij}}}{\boldsymbol{K_{i-1,j}}}=1\left|\frac{\boldsymbol{I_{i-1,j}}}{\boldsymbol{K_{i-2,j}}},\boldsymbol{z}_j,\boldsymbol{x}_{ij},\boldsymbol{\zeta}_j\right.\right)\right\}=\gamma_{z_0}+\alpha\left(\frac{\boldsymbol{I_{i-1,j}}}{\boldsymbol{K_{i-2,j}}}\right)+\gamma_{x_1}\left(\frac{\boldsymbol{S_{ij}}}{\boldsymbol{k_{i-1,j}}}\right)+\gamma_{x_2}\left(\mathbf{Emp}_{\boldsymbol{ij}}\right)+\boldsymbol{\zeta_j}$$

where $\boldsymbol{x_{ij}} \in \left[\frac{\boldsymbol{S_{ij}}}{\boldsymbol{k_{i-1,j}}}, \text{Emp}_{\boldsymbol{i-1,j}}\right]$ and there are no time-invariant, $\boldsymbol{z_j}$, covariates. In this setting, $\frac{\boldsymbol{I_{ij}}}{\boldsymbol{K_{i-1,j}}}$ is still a binary response variable taking the value of 1 if firm j invests at occasion i and 0 otherwise. The associated component of the joint model is as before where the initial response is modelled at $i = 0$

$$\text{probit}\left\{\left(\frac{\boldsymbol{I_{0j}}}{\boldsymbol{K_{0,j}}}=1\left|\boldsymbol{z}_j,\boldsymbol{x}_{0j},\boldsymbol{\zeta}_j\right.\right)\right\}=g_{z_0}+g_{x_1}\left(\frac{\boldsymbol{S_{ij}}}{\boldsymbol{k_{i-1,j}}}\right)+g_{x_2}\left(\mathbf{Emp}_{\boldsymbol{ij}}\right)+\alpha\left(\frac{\boldsymbol{I_{i-1,j}}}{\boldsymbol{K_{i-2,j}}}\right)+\lambda_0\boldsymbol{\zeta_j}.$$

The empirical auxiliary model of within-means is constructed as follows

$$\boldsymbol{\zeta_j}=\delta_{\bar{x}_1}\left(\frac{\widetilde{\boldsymbol{s_j}}}{\boldsymbol{k_{1,j}}}\right)+\delta_{\bar{x}_2}\left(\widetilde{\boldsymbol{Emp}}_{\boldsymbol{2,j}}\right)+u_j.$$

In the case of the conditional model, we implement the following auxiliary model

$$\boldsymbol{\zeta_j}=\delta_{y_{0j}}\left(\frac{\boldsymbol{I_{0j}}}{\boldsymbol{K_{0,j}}}\right)+\delta_{x_{10}}\left(\frac{\boldsymbol{s_{1j}}}{\boldsymbol{k_{1,j}}}\right)+\delta_{x_{20}}\left(\mathbf{Emp}_{\boldsymbol{2,j}}\right)+\delta_{\bar{x}_1}\left(\frac{\overline{\boldsymbol{s_{1j}}}}{\boldsymbol{k_{1,j}}}\right)+\delta_{\bar{x}_2}\left(\overline{\mathbf{Emp}_{\boldsymbol{2,j}}}\right)+u_j.$$

5.2.2 Patterns of Investment Decisions and Estimates of the Structural Investment Model

The descriptive analysis covered in this section presents low patterns of participation of firms in capital investments. With missing data, it is possible to analyse all survey waves for which the investment rate y_{ij} and associated explanatory variables x_{ij} are not missing for a subject. It is also useful to consider each occasion that precedes an occasion with missing data as an initial occasion and assume that the second line of Eq. 9 holds for all initial responses. As in Hyslop (1999) and Chay and Hyslop (2000), in order to improve our understanding of the fit of the models estimated, we first present frequencies of a firm's discrete choice to invest in a given occasion as shown in Table 8.

For each sequence in Table 8, a "1" in the i^{th} position denotes an observed positive investment in the i^{th} period, whereas a "0" indicates a missing value of investment. For example, the pattern of missingness characterized by the sequence '0000000000' in Panel A indicates that 100 out of 227 firms have no responses for investment in any of the 10 years from 1994-2003, while '0111111111' in Panel C means only one out of the same number of firms invested consecutively after the first year of inaction in the sample. However, isolated observations that follow sequences like '0101010101' cannot be used because only initial values are supplied rather than the required consecutive sequences. Nonetheless, several sequence types of non-missing values of investment participation by a firm can be used, e.g. '1101100100'. In this case, the initial response is y_{0j} for the first sequence and y_{3j} for the second sequence and so on. The parameters of the auxiliary model can then vary according to the location of the initial occasion. Another practical matter is to analyse only contiguous sequences of non-missing data that start at occasion 0 and discard firms with patterns of the form '0101000000'. In such *ad hoc* approaches, the missing values of x_{ij} are implicitly imputed by $\overline{x}_{.j}$ and y_{ij} is assumed to be missing at random (MAR).

Analysing relationships between the response variable and covariates based on either contiguous investments or investments with non-missing patterns ensures accurate estimation of the likelihood function and unbiased parameter estimates, see Skrondal and Rabe-Hesketh (2014) and Seaman Galati, Jackson and Carlin (2013)[21]. However, this might present us with the technical problem of 'not enough observations' prevalent in finite samples with short *T*, see Akay (2012) and Albarran *et al*. (2015).

[21] See Seaman *et al*. (2013) on handling "Missing at Random" and "Missing Completely at Random" datasets as well as potential implications for the likelihood function.

Table 8: Manufacturing Patterns of Missing Values and Investment Participation $\left(\boldsymbol{y_{ij}} = \frac{I_{ij}}{K_{ij-1}}\right)$ in Swaziland (1994-2003)

Panel A				Panel B				Panel C			
Missing Values' Patterns	Freq.	Percent	Cum.	Missing Values' Patterns	Freq.	Percent	Cum.	Missing Values' Patterns	Freq.	Percent	Cum.
0000000000	100	44.05	44.05	0000101000	1	0.44	72.25	0011101111	1	0.44	87.22
0000000001	10	4.41	48.46	0000101111	1	0.44	72.69	0011111010	1	0.44	87.67
0000000010	7	3.08	51.54	0000110001	1	0.44	73.13	0011111101	1	0.44	88.11
0000000011	5	2.20	53.74	0000110010	1	0.44	73.57	0011111111	2	0.88	88.99
0000000100	4	1.76	55.51	0000111000	2	0.88	74.45	0100000000	3	1.32	90.31
0000000101	1	0.44	55.95	0000111110	2	0.88	75.33	0100001110	1	0.44	90.75
0000000110	2	0.88	56.83	0000111111	1	0.44	75.77	0100011110	1	0.44	91.19
0000000111	4	1.76	58.59	0001000000	2	0.88	76.65	0100011111	2	0.88	92.07
0000001000	4	1.76	60.35	0001001100	2	0.88	77.53	0100100000	1	0.44	92.51
0000001001	2	0.88	61.23	0001001110	2	0.88	78.41	0100100011	1	0.44	92.95
0000001010	2	0.88	62.11	0001010111	1	0.44	78.85	0100101110	1	0.44	93.39
0000001011	1	0.44	62.56	0001011110	2	0.88	79.74	0100111110	1	0.44	93.83
0000001100	1	0.44	63.00	0001100000	2	0.88	80.62	0101111101	1	0.44	94.27
0000001110	5	2.20	65.20	0001100100	1	0.44	81.06	0110000000	1	0.44	94.71
0000001111	2	0.88	66.08	0001101111	1	0.44	81.50	0110000010	1	0.44	95.15
0000010000	2	0.88	66.96	0001111011	1	0.44	81.94	0110000111	1	0.44	95.59
0000010001	1	0.44	67.40	0001111110	2	0.88	82.82	0110101000	1	0.44	96.04
0000010110	1	0.44	67.84	0001111111	1	0.44	83.26	0110110000	1	0.44	96.48
0000011000	1	0.44	68.28	0010000001	1	0.44	83.70	0110111000	1	0.44	96.92
0000011001	1	0.44	68.72	0010000011	1	0.44	84.14	0110111111	1	0.44	97.36
0000011011	1	0.44	69.16	0010011010	1	0.44	84.58	0111000000	2	0.88	98.24
0000011100	1	0.44	69.60	0010011111	1	0.44	85.02	0111000110	1	0.44	98.68
0000011110	2	0.88	70.48	0010100000	1	0.44	85.46	0111110000	1	0.44	99.12
0000100000	1	0.44	70.93	0010100100	1	0.44	85.90	0111111110	1	0.44	99.56
0000100010	1	0.44	71.37	0010100111	1	0.44	86.34	0111111111	1	0.44	**100.00**
0000100011	1	0.44	71.81	0011000000	1	0.44	86.78	**Total**	**227**	**100.00**	

A total of 82 out of 227 observations in the sample have at least 2 consecutive non-missing sequences, implying that only 36.12 percent of the firms provide descriptive evidence of some serial persistence. Viewed with the high incidence of inaction, this suggests the possibility that the underlying process is largely independent over time. These investment transitions point to adopting a model that includes: a first-order Markov chain to capture any degree of true state dependence, and/or serially correlated errors as well as unobserved heterogeneity in order to fit the sequences, see Rabe-Hesketh and Skrondal (2014).

Table 9: Multilevel Parameter Estimates and Robust Standard Errors for Dynamic Random Effects Probit Models of Investment.

Structural Parameters	Naïve	Estimates for Joint Models: Exogenous x_{ij}	Endogenous x_{ij}	Estimates for Conditional Maximum Likelihood Model: Conditional Estimator	NPMLE (*Mass Point Method*)
$\left(\frac{I_{i-1.j}}{K_{i-2.j}}\right)$	2.0634**	0.2189	0.5556	0.9143	0.7327
	(0.6817)	(0.7428)	(0.6900)	(0.6566)	(0.6166)
$\left(\frac{S_{ij}}{k_{i-1.j}}\right)$	−0.7291	0.2655	0.1574	1.0331	1.2758
	(0.3892)	(0.5596)	(0.4487)	(0.7938)	(0.7857)
(Emp_{ij})	0.2341*	0.0928	0.0526	0.2921	0.7240
	(0.1003)	(0.1054)	(0.4742)	(0.4443)	(0.5415)
Nolag		−5.5564***	−4.9138**		
		(1.5658)	(1.5898)		
$\left(\frac{S_{ij}}{k_{i-1.j}}\right) X$ **Nolag**		1.8706	2.0390		
		(1.1988)	(1.1943)		
$(\mathrm{Emp}_{ij}) X$ **Nolag**		0.5516**	0.4258*		
		(0.2081)	(0.2084)		
$\left(\frac{S_{ij}}{k_{i-1.j}}\right)^{O}$				−2.8371*	−2.5949
				(1.2537)	(1.4484)
$(\mathrm{Emp}_{ij})^{O}$				−0.7460	−1.3585
				(0.5493)	(0.7236)
$\left(\overline{\frac{S_{.j}}{k_{1.j}}}\right)$				0.7728	0.1701
				(1.0795)	(0.7960)
$\left(\overline{\mathrm{Emp}_{2.j}}\right)$				0.8757	1.2089
				(0.6294)	(0.7561)
Constant	1.0008	1.5484*	0.7161	0.6352	0.9435
	(0.6159)	(0.7167)	(0.6804)	(0.9313)	(1.6922)
cbri1					
ψ	0			0.3770*	
	0			(0.1761)	
cbr1_1l					
Nolag		−0.2818	0.1866		
		(0.3313)	(0.6448)		
cbr1_1					
One		0.8661***	0.5763**		
		(0.2040)	(0.2081)		
f1:					
$\left(\widetilde{\frac{S_{.j}}{k_{1.j}}}\right)$			−0.139		
			(0.2169)		
$\left(\widetilde{Emp}_{2.j}\right)$			0.1995		
			(0.4335)		
z2_1_1					
Constant					−0.9755
					(1.4523)
p2_1					
Constant					0.8836
					(0.4512)
Number of Firms	350	911	626	480	480
Log-likelihood	−95.3007	−184.086	−166.715	−133.386	−130.03

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Table 9 presents estimates of the empirical model and distinguishes each model on the basis of various assumptions about the initial conditions problem and endogeneity of covariates. The Naïve results are presented in Model 1. The joint distribution model coefficients are presented in Model 2 and Model 3. Model 4 presents results for the conditional model while Model 5 presents results for the same model using the NPMLE methods based on the mass point procedure.

The Naïve specification uses all the available observations for the dependent variable. The model is estimated with Stata's **xtprobit** command to produce biased standard errors, but correct parameters (see Skrondal and Rabe-Hesketh, 2014). The routine achieves this by performing a sensitivity evaluation of the results using quadrature checks, and we keep adding an integration point until the log-likelihood remains unchanged. However, since the standard errors are biased upward, we also use the **gllamm** command and adaptive quadrature for accurate point estimates and robust standard errors. In this model, as shown in column (1), longitudinal dependence is almost completely due to state dependence as a result of ignoring initial conditions and endogeneity of covariates. As expected, the coefficient of investment rate at $t-1$ is significant and large at 2.06 percent, spuriously suggesting significant persistence of true state dependence of investment rates. The estimated variance of the random-intercept is 0.00.

In estimating the joint distribution model with *exogeneity* assumption, all available data, including the missing investment rate lag, are used. The approach adopted allows for different coefficients for initial responses. Although still positive, the coefficient on lagged investment rate is greatly reduced in absolute terms to 0.22 percent and is insignificant at conventional levels. A dummy variable, **Nolag**, represents all observations with missing data on investment rates at time $t-1$ and enters the model significantly at 1 percent level. It is not surprising that the dummy is negative and significant, given the high incidence of single investment rates that are sandwiched between missing values in Swazi manufacturing reflected in Table 4.8. This means a unit percentage point increase in net PME investment inactivity at $t-1$ reduces the probability of investment by $[-5.56, -4.91]$ percent at time t. When **Nolag** is interacted with $\mathbf{Emp}_{ij}$, it produces a positive and significant coefficient at the 10 percent level. This is consistent with larger firms, measured in terms of employment size, not investing at $t-1$. The larger firms' reasons for this might be related to the potential substitution of capital adjustment plans for increased (possibly fixed contract) labour at time t.[22] Such decisions would continue until the uncertainty about the Southern African economic outlook brought about by trade reforms in the 1990s was resolved, see similar arguments by Bloom (2009) for the U.S. case.

The *endogeneity* assumption concerning covariates in the joint distribution model also uses all available data. However, in contrast to the exogeneity model, it relies on *different* coefficients for all initial responses. Its longitudinal means needed to obtain an appropriate linear predictor for consistent estimation are based on occasions where the investment rate variable is not missing. Notably, the estimated factor loading for the linear predictor multiplying the random-intercept enters the auxiliary

[22] Capital irreversibility and the relative ease of employment termination for contract workers are assumed.

model insignificantly at all conventional levels.[23] Furthermore, Skrondal and Rabe-Hesketh (2014) suggest that a test of $H_0: \bar{x}_{\left(\frac{\widetilde{s_{.j}}}{k_{1.j}}\right)} = \bar{x}_{\left(\widetilde{Emp}_{2.j}\right)} = 0$ is equivalent to a level 2 test of exogeneity. Since both statistics are insignificant, the exogeneity hypothesis cannot be rejected. Thus, there is no material difference between the results of the two joint distribution models and therefore the key predictions of the model under exogeneity assumptions are maintained.

An alternative to the joint distribution models used in this analysis is the conditional model that conditions on initial responses and explanatory covariates. Instead of using all available data, the method is designed to rely on consecutive sequences of *at least* two non-missing values of investment rates in order to analyse contiguous sequences only. However, the change in the definition of the response variable poses a barrier to the estimation of the model when the dataset is densely populkated with missing values in the investment series as shown in the descriptive analysis of Section 4.2.[24] When erratic investments are excluded in the analysis, only about two firms have at least two consecutive sequences of investment in each of the patterns of missing values during the ten-year period.

As a consequence of this difficulty, the estimation of the conditional model is now based on all the data and the coefficients for initial period explanatory variables are unfortunately constrained to be equal to coefficients of subsequent periods. For a critique on using the entire sample and initial conditions to compute within-firm means, see Rabe-Hesketh and Skrondal (2013). The model is therefore estimated just to provide upper bounds for coefficients of the joint distribution models. Thus, the estimated random-error variance is 0.377 and the associated intraclass correlation of the latent variable, y_{ij}^*, in Eq. 9, given the observed sales/capital ratio and labour, is $\frac{\psi}{\psi+1} = 0.27$. That is, approximately 27 percent of the variance in real investment rates that is not explained by the observed covariates is produced by unobserved time-invariant firm-specific characteristics. Similarly, the suitability of the restricted one-factor model is measured by the statistical insignificance from unity of λ_i, that is, $\lambda_i = 1$. This is estimated to range between $[0.58,\ 0.87]$.

We also use the NPMLE approach to replicate the conditional model results by using the Rabe-Hesketh *et al.* (2005) adaptive quadrature to maximize the likelihood function and determine the optimal mass-point

[23] This indicates that the random-intercept regressed on longitudinal means based on non-missing investment rates in the auxiliary equation, $\left(\frac{\widetilde{s_{.j}}}{k_{1.j}}\right)$ and $\left(\widetilde{\boldsymbol{Emp}}_{\mathbf{2}.j}\right)$, can be used in the generalized linear latent and mixed modelling approach embedded in **f1: a. b**, where **a** and **b** represent a one-factor probit model described in Arulampalam and Stewart (2009)**.** These are averages representing the extent to which item i, in an item response setting, discriminates between firms of different propensities to invest thereby allowing the analyst to extract unobserved heterogeneity.

[24] An experiment conducted using the Heckman (1981a) estimator for serially independent idiosyncratic shocks using Stewart (2006) failed to estimate the probit model for $t = 1$ due to insufficient observations.

based on the Gâteaux derivative method. This technique avoids making any assumptions about the distribution of the random-intercept; see Heckman and Singer (1984) and Rabe-Hesketh *et al*. (2003) for details on this method. It produces structural coefficients that are similar to those of the conditional model and are indeed systematically greater in absolute terms than those produced by the joint models.

5.2.3 Estimation of an Endogenous Regime Switching Model

The analysis thus far has focussed largely on the properties and estimation of true state dependence as well as individual firm-specific heterogeneity underlying firms' investment choices. It is of interest therefore to also study the parametric patterns of the proxy of marginal q to determine if there is any switching of investments across different regimes as implied by Abel and Eberly (1994) and recently Abel (2014). Firms may sort their investments in terms of either high or low regimes as in Drakos and Konstantinou (2013) for the case of Greece. The next sections concentrate on this task.

This section presents estimates of the structural investment equation using full information maximum likelihood (FIML) methods. This efficient method for estimating the endogenous regime switching regression model was first proposed by Lee and Frost (1978), and described for Stata by Lokshin and Sajaia (2004). Such methods simultaneously estimate the discrete probit criterion or selection equation and the continuous model to produce consistent standard errors. It sorts out investment rates according to two different states and simultaneously estimates the binary and continuous components of the empirical model. Firstly, the two sets of parameters of interest are β^H and β^L representing the high and low investment regimes respectively, which measure the effects of the $t-1$ covariates that determine investment rates. The second parameter vector is γ which measures the effects of $t-1$ covariates included in the switching function. Thirdly, the standard deviations of ε_{it}^H and ε_{it}^L; namely, $\sigma_{H\varepsilon}$ and $\sigma_{L\varepsilon}$, can be estimated. Lastly, the correlation coefficients $\rho_{H\mu}$ and $\rho_{L\mu}$ in both investment regimes are easy to estimate. Thus, the endogenous switching regression model of investment is suitable for estimating this model.

It is common practice in selection models like ours to introduce a variable(s) that can produce nontrivial variation in the selection part of the model while not affecting the outcome variable directly. Although three variables; namely, material input, energy and the inverse of firm-size measure are available, the latter is adopted here because it affects only the extensive margin of investment in the switching function rather than the intensive margin (see Letterie and Pfann, 2007).[25] This implies three scenarios: (1) that if larger firms are more likely to locate in the higher investment regime, the firm-size

[25] However, the exclusion restriction may cause global concavity failure in some settings, in which case the model may be identified by nonlinearities thereby causing the selection equation to contain only the regressors in the continuous equations.

measure, $(K_{it})^{-1}$, will produce a negative sign. (2) In contrast, the sign will be positive if smaller firms have a higher propensity to locate in the high investment regime. (3) If firm selection into the high investment regime is scale-independent, then the exclusion restriction imposed by the introduction of the inverse of capital stock will be insignificantly different from zero in the switching function.

Table 10 presents results of a structural endogenous switching regression model which reveal some form of existence and differences in high and low investment expenditures in PME at the firm level in Swaziland.[26] In order to make inferences about investment behaviour between regimes, two tests are conducted principally for Model 1 and Model 4 because of their central role in the GMM approach in the previous section. This exercise is also performed for the other components of fundamentals; that is, the squares, averages and squares of averages of each model to determine their individual behaviour across regimes.

In the case of the investment response to movements in the $t-1$ employment and its components, we find insignificant coefficients in the high regime and highly significant and negative coefficients in the low regime. More specifically, firms in the high investment regime category of Models 1-3 substitute investment expenditure in PME for employment insignificantly while low investment regime firms chose a relatively higher capital-labour substitution pattern. In all three cases, the single regime hypothesis $H_0 \colon \beta^{High} = \beta^{Low}$ is not supported by the χ^2–distribution of the Wald-test statistic at the 1 percent level.[27] For example, this is χ^2 (1) = 269.67 with *p-value*=0.0000 for the linear relationship expressed in Model 1. In this model, given the strong empirical evidence that the data generating process is consistent with two significantly different regimes, it is instructive to discuss the variables influencing the likelihood that an observation belongs to the high or low investment regime. Since the coefficient of the investors of the capital stock, $(K_{it})^{-1}$, is insignificant, the location of an observation in the high regime is not a function of firm size. Therefore, the endogenous switching regression results confirm visually and technically that the dataset is generated by two investment regimes in the capital adjustment-employment nexus, in contrast to the single regime structure presented through the systems−GMM approach.

[26] Convergence difficulties of the likelihood function, even after changing starting values, required a slight adjustment in the presentation of the empirical model results, in contrast to Letterie and Pfann (2007). This allowed us to analyse each component of the structural model separately as Hu and Schiantarelli (1998) for the U.S. case.

[27] Hu and Schiantarelli (1998), Nielsen and Schiantarelli (2003) and Letterie and Pfann (2007) note the difficulty computing the degrees of freedom if the null hypothesis holds because the parameters in the switching function are unidentified, and the likelihood ratio (LR) test might not even have a χ^2–distribution. Goldfeld and Quandt (1973) also show that the use of a χ^2–distribution for the LR−test with degrees of freedom equal to the number of constraints *plus* the number of unidentified parameters yields a test that favours non-rejection of the restrictions.

Table 10: FIML Estimation of Endogenous Switching Regression Models: 1994-2003

$Z_{it}=\left\{Emp_{t-1}, Emp_{t-1}^2, \overline{Emp}, \frac{S_{it-1}}{K_{it-2}}, \frac{S_{it-2}}{K_{it-3}}, \left(\frac{\overline{S}}{K}\right)^2, (K_{t-1})^{-1}, \text{YD}, \text{ID}\right\}$

Regime	Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$\left[\frac{I_{it}}{K_{it-1}}\right]^{High}$	Emp_{t-1}	−0.132					
		(0.0776)					
	Emp_{t-1}^2		−0.009				
			(0.0086)				
	$\overline{Emp}$			−0.106			
				(0.0819)			
	$\frac{S_{it-1}}{K_{it-2}}$				−0.001*		
					(0.0003)		
	$\left(\frac{\overline{S}}{K}\right)$					−0.003***	
						(0.0005)	
	$\left(\frac{\overline{S}}{K}\right)^2$						0.000***
							(0.0000)
	Constant	7.803***	7.396***	7.715***	7.203***	7.208***	7.186***
		(0.368)	(0.2298)	(0.3942)	(0.4299)	(0.333)	(0.3276)
$\left[\frac{I_{it}}{K_{it-1}}\right]^{Low}$	Emp_{t-1}	−0.821***					
		(0.1829)					
	Emp_{t-1}^2		−0.110***				
			(0.023)				
	$\overline{Emp}$			−0.962***			
				(0.1956)			
	$\frac{S_{it-1}}{K_{it-2}}$				−0.451**		
					(0.1652)		
	$\left(\frac{\overline{S}}{K}\right)$					−0.439***	
						(0.092)	
	$\left(\frac{\overline{S}}{K}\right)^2$						−0.026***
							(0.0043)
	Constant	5.854***	4.553***	6.058***	3.373	4.281***	3.904***
		(0.4943)	(0.4837)	(0.497)	(3.0102)	(1.089)	(1.0567)
Switching Function	Emp_{t-1}	0.386***					
		(0.068)					
	Emp_{t-1}^2		0.049***				
			(0.0096)				
	$\overline{Emp}$			0.398***			
				(0.068)			
	$\frac{S_{it-1}}{K_{it-2}}$				0.076		
					(0.0618)		
	$\left(\frac{\overline{S}}{K}\right)$					0.040*	
						(0.0176)	
	$\left(\frac{\overline{S}}{K}\right)^2$						0.002
							(0.005)
	K^{-1}	−0.154	−0.187	−0.073	0.785	1.384	1.846
		(0.1491)	(0.1673)	(0.3839)	(3.1523)	(1.1185)	(1.301)
	Constant	−0.147	0.566	−0.523*	0.479	0.568	0.56
		(0.4144)	(0.3813)	(0.2622)	(0.5853)	(0.4119)	(0.4331)
Statistics	$\sigma_{H\varepsilon}$	1.8183	1.7881	1.8476	1.6548	1.7930	1.7846
		(0.0947)	(0.0958)	(0.1003)	(0.4094)	(0.3629)	(0.3649)
	$\sigma_{L\varepsilon}$	1.3066	1.2785	1.3626	1.6008	1.3316	1.2839
		(0.2412)	(0.2265)	(0.2534)	(1.4987)	(0.2741)	(0.1474)
	$\hat{\rho}_{H\mu}$	−0.9167	−0.8810	−0.9066	−0.8816	−0.8833	−0.874
		(0.06)	(0.0819)	(0.0631)	(0.4736)	(0.4027)	(0.4304)

$\hat{\rho}_{L\mu}$	−0.6075	−0.5779	−0.6723	−0.7911	−0.48047	−0.2744
	(0.2886)	(0.2976)	(0.2270)	(0.9093)	(0.6104)	(0.6129)
NT	378	378	378	252	358	358
Log Likelihood	−820.14	−820.188	−830.692	−540.799	−790.234	−791.973

H_0: $\beta^{High} = \beta^{Low}$ for Model 1: χ^2 (1) = 269.67, Prob > χ^2 = 0.0000

H_0: $\beta^{High} = \beta^{Low}$ for Model 4: χ^2 (1) = 0.75, Prob > χ^2 = 0.3868

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

The adopted proxy for the shadow price of capital and related components in Models 3-4 produced significant and negative results in both regimes, but more so in the low investment regime. That is, the coefficients for the sale/capital ratio in the high and low investment regimes are $\beta^{High} = -0.001$ and $\beta^{Low} = -0.439$, respectively. The standard levels of significance suggest that $\beta^{High} \neq \beta^{Low}$ in the statistical sense. However, the χ^2–distribution of the Wald-test statistic supports the equality null hypothesis for Model 4 coefficients at χ^2 (1) = 0.75 with *p-value*=0.3868. These results are robust to chosen transformations of the proxy variable for the shadow price of capital. This means that the investment function can be expressed as a single investment regime problem and therefore the parameters of the switching function are not identified and validates the conclusions drawn from the systems−GMM approach.

Finally, the correlation coefficients, $\hat{\rho}_{H\mu}$ and $\hat{\rho}_{L\mu}$, measure the relationships between the error terms in the high and low investment regimes and the error term in the switching function. As in Nielson and Schiantarelli (2003, footnote 26) and Letterie and Pfann (2007, p. 810), the statistic $\hat{\rho}_{H\mu} \to 1$ in absolute terms, which is typical of switching models.[28] Our results mimic those of Hu and Schiantarelli (1998) for the U.S., Nielsen and Schiantarelli (2003) for Norway and Letterie and Pfann (2007) for the Netherlands.

6 Discussion of Results

In the analysis of the dynamic structural model of investment, the descriptive statistics show patterns of significant microeconomic lumpiness and discontinuous investment in plant, machinery and equipment (PME). The data is characterized by a high incidence of zero investment rates and this stylized fact is distribution-free. Even if the data is divided into investments with or without expenditure on maintenance and repair (M&R), it still produces a high incidence of zeros at 44 percent and 73 percent, respectively. For ease of comparison with other country studies, the analysis subsequently focuses on the data with investment cost of M&R. As a result, only 36.12 percent of observations have a sequence of at least two consecutive non-missing values of investment. Considering the ten-year span of investment inactivity for a significant number of establishments in

[28] See Goldfeld and Quandt, (1973). Hu and Schiantarelli (1998) break their sample into two samples to minimize endogeneity problems induced by the correlation between the error terms in the investment functions and the switching equation. However, this creates new problems by imposing restrictions on the nature of the firm-specific effects.

manufacturing, the sector was characterized by deepening capital obsolescence and a potential decline in capital productivity. Investments also feature a mesokurtic; that is, skewness and high kurtosis in investment distribution. These preliminary empirical regularities already suggest that the microeconomic industrial capital adjustment costs in Swaziland are nonconvex, and can translate to similar aggregate patterns as in Cooper *et al*. (1995) and Khan and Thomas (2008).

Looking at industrial investment hazard functions, slicing the data into groups of small and large plants produces interesting results. A firm's discrete choice to invest in PME appears to be scale dependent. Large firms' propensity to invest in excess of 20 percent is significantly higher than that of small firms at standard statistical levels. This story remains unchanged when the definition of an investment spike is reduced to 10 percent. However, the probability of an industrial spike for either group of plants is less than seven percent during the period under study. This re-enforces the earlier conclusion about general investment passivity among Swazi firms during the entire period of trade liberalization. Our conjecture is that this period ushered in new market competition that forced inefficient establishments out of business while foreign plants relocated back in home markets to experience economies of scale. For remaining firms, the re-integration of South Africa back to the world economy brought substantial business uncertainty in the customs union which required Swazi firms to monitor their own market share dynamics and hold back on major new capital investments.

The data is further taken to rigorous analysis using a structural model of investment to establish the impact of state dependence of investment decisions and the sales/capital ratio, controlling for plant size. We begin with generalized method of moments (GMM) estimators that exploit orthogonality conditions applied to the theoretical model. This effort produces imprecise coefficients of previous investment, sales/capital ratio and employment. One obvious source of imprecision in the estimation of model parameters is the small sample size of firms and high investment heterogeneity. This obscures any potential persistence in the investment rate series. Another likely explanation involves omitted variables that may be correlated with included regressors and this has confounding effects on parameters. Nonetheless, the orders of magnitude and the signs of the coefficients remain consistent with findings in the larger literature; see for example, Drakos and Konstantinou (2013) for the case of Greece.

This approach is subsequently extended to a multilevel discrete choice binary data analysis that allows for both longitudinal within-firm dependence and unobserved heterogeneity. It further makes provision for the direct analysis of the impact of a firm's option to exercise its option to wait-and-see in an uncertain environment and the associated interaction with sales/capital ratio and firm-size. The method we use to handle initial conditions and endogenous explanatory variables in the model of binary data with unobserved heterogeneity confirms the GMM results. That is, the theoretical model's parameters are still imprecisely measured. However, when firms defer investment by a single lag, the cost of exercising this option is 5.56 percent or 4.91 percent depending on whether covariates are assumed exogenous or endogenous, respectively. Although interacting investment lags with the sales/capital ratio remains insignificant, the results change when the $t-1$ investment lag is interacted with firm-size. When a large firm begins investment after a single period of inactivity, this increases investments by 0.55 percent or 0.42 percent in the manufacturing sector depending on exogeneity-endogeneity assumptions made about the covariates, respectively.

The scale-dependence of industrial investment patterns in Swaziland is further subjected to a framework that provides for endogenous switching of firms between high and low investment regimes, β^H and β^L; respectively. This helps in our understanding of the behaviour of investment patterns by small and large firms in periods of uncertainty in a small member of a customs union. This procedure reports a valid switch of investments between the high and low regimes. The high regime coefficients, β^H, are either insignificant or persistently zero but negatively charged. On the other hand, the low investment regime switchers report significantly negative coefficients, β^L. However, these results fail to corroborate the validity of scale-dependence in investments if the definition of firm size changes from employment to the inverse of real capital stock as in Letterie and Pfan (2007). That is, $(K_{it})^{-1}$ is insignificant.

7 Summary and Conclusion

This paper investigates the presence of state dependence, unobserved heterogeneity and the impact of real sales/capital ratio on investment rates for the manufacturing sector in Swaziland. It begins with a descriptive analysis of a panel dataset for 13 industries and finds that the rate of investment is as low as 0.24 percent every year, with the observed investment heterogeneity measured by the standard deviation just as low at 0.29.

The analysis of the microeconomic investment spike hazard confirms the lacklustre investment patterns of the manufacturing sector during the period of trade liberalization in the Customs Union. Using the definition of an investment spike presented by Cooper *et al.* (1999), which is investment rate in excess of 20 percent, the probability distribution of spiky events is less than 0.07. The fact that the empirical hazard is upward-sloping is taken as evidence of within-plant effects instead of between-plant effects. These lumpy investments are scale-dependent in that they are dominated by firms employing more than 50 workers.

Using a structural model of investment, we investigated the effects of past investments, unobserved heterogeneity and real sales/capital ratio on investment rates by relying on three methods. These were the GMM approach, multilevel random-effects and the switching regression regime methods. We find that the impact of true state dependence is insignificant in all models. That is, previous investment has no influence on the current decision to invest in the Swazi manufacturing sector. However, this is not an indictment of the conventional wisdom that the best predictor of current investment is its lagged levels as discussed in the descriptive analysis section. It is simply a reflection of a non-investing sector because of high uncertainty and associated firm-level entry/exit dynamics. The ratio of real sales/capital also has insignificant effects on investment rates across model specifications. Furthermore, the impact of unobserved individual firm characteristics underlying the discrete choice to invest in PME had insignificant effects on investment rates. This suggests that technological change does not translate into transformational and entrepreneurial investment. Therefore, unobserved heterogeneity and investment dynamics confer insignificant effects in the structural model.

However, allowing for self-selection of firms into high and low investment regimes, the endogenous regime switching model adds more clarity on the results obtained in the GMM and multilevel random intercept models. While the empirical hazard function displays large firm dominance in spiky investment episodes, the endogenous investment regime switching model produces scale-independent results. That is, the inverse of capital stock remains insignificant in all model specifications. Thus, both firm sizes locate in either regime in the manufacturing sector. More specifically, high regime investors are largely in the zone of inaction while those in the low investment regime are accountable for observed disinvestments. A Wald-test of model independence in the switching regression model is generally confirmed. Firms are subject to common exogenous shocks of trade liberalization, and investment decisions are characterized by herd behaviour leading to a dominant response of exercising the option to withholding $t-1$ investment until period t.

Our structural model only included time-variant covariates. The variation in investment rates that is not explained by changes in marginal q, investment dynamics and employment is captured through the intraclass correlation of 27 percent. That is, omitted time-constant regressors might also be important in the model.

For the first time as far as we are aware, we obtain the most interesting results when the impact of missing values of *net* PME expenditure on the investment rate is investigated in more depth. This involves identifying firm-level consecutive sequences of positive investments, along with instances of non-response to capture cases with no $t-1$ investment values. The impact of such missing values significantly reduces the decision to invest by $[4.91,\ 5.56]$ percent. This means that the cost of delaying investment in the sector by one period is a reduction of investment in the next period by a significant percentage in Swaziland. An increase in the lag depth of firms' exercising of the option to wait before investing generates an increase in the industrial investment cost. However, when the incidence of missing values is interacted with employment, the probability of investment substitution by employment is significantly increased by $[0.43,\ 0.55\ ]$ percent in the sector. This means that the lack of robust investment in capital goods in the sector was compensated for by increasing employment, *at the margin*.

As a whole, this means that the manufacturing sector in Swaziland experienced a high incidence of zero investment in plant, machinery and equipment in the period 1994-2003. More specifically, firms refrained from large capital investments for the establishment of new plants but maintained and repaired the machinery and equipment to keep business operations running. This was complemented with some capital substitution for employment. A consequence of this investment behaviour by Swazi manufacturers was a deterioration and becoming obsolete of capital goods in the sector, leading to a general decline in technological advancement. Another effect involved the loss of predictive power of current investment concerning future investments. The timidity and herd behaviour observed among firms also dampened any impact of unobserved industrial heterogeneity. That is, firms' capital adjustment plans were largely similar among producers and insignificant across industries.

Since the conditional model in this paper conditions on initial responses and explanatory covariates, and uses consecutive sequences of *at least* two non-missing values of investment rates to analyse contiguous sequences, it fails in datasets with limited successive sequences. Therefore, the next research agenda involves nonlinear methods of estimation that directly account for the unbalanced nature of investment data. Such methods need to allow for the use of all available observations while relaxing the assumption that observations are completely missing at random. A similar idea is conceptualized by Albarran *et al.* (2015) who develop some dynamic nonlinear random effects models with unbalanced panels based on all available information. Wooldridge (2010) also presents useful correlated random effects models with unbalanced panels.

The structural model studied here can be extended in other directions. One such extension would involve relaxing the first-order Markov structure by considering an increased lag depth of the investment rate or by specifying models where the lagged investment has time-varying parameters, see Skrondal and Rabe-Hesketh (2014). Second, Francis, Stott and Davies (1996) and Albert and Follmann (2003) construct models which allow covariate parameters and the impact of the random intercept to depend on own previous states. Third, a direct extension of this work can also entail nominal, ordinal or censored responses or counts, including the conditional approach discussed in Wooldridge (2005) for various response types. Fourth, investment dynamics can be expressed in terms of latent Markov models; that is, in terms of $y^*_{i-1,j}$ as in Pudney (2008). Alternatively, we could follow Heckman (1981a) who generalizes a transition model of binary responses that incorporates lags for both observed and latent responses. Fifth, we could relax the longitudinal independence assumption concerning the level 1 error as in Hyslop (1999), Stewart (2006) and Hajivassiliou and Ioannides (2007). Sixth, the use of a random intercept to specify unobserved heterogeneity could be replaced with more general specifications involving several random coefficients or common factors as in Heckman (1981a), Bollen and Curran (2004) and Skrondal and Rabe-Hesketh (2014).

APPENDICES

Appendix A4.1: Definition of Terms for the GMM Estimation

Endogeneity: when x_{it} is endogenous, it is correlated with current and deeper lags of shocks; i.e. $E(\varepsilon_{it}|x_{is}) = 0 \; \forall \; t > s$ and $E(\varepsilon_{it}|x_{is}) \neq 0 \; \forall \; t \leq s, \forall \; i = 1, \dots, N, \forall \; t, s = 1, \dots, T$. Lagged values dated $t{-}2$ and earlier are therefore valid instruments; hence, variables in first differences are instrumented with suitable lags of their own levels.

Predetermined: when x_{it} is predetermined, it is uncorrelated with future shocks but is correlated with their lags; i.e. $E(\varepsilon_{it}|x_{is}) = 0 \; \forall \; t \geq s$ and $E(\varepsilon_{it}|x_{is}) \neq 0 \; \forall \; t < s, \forall \; i = 1, \dots, N, \forall \; t, s = 1, \dots, T$. The first differenced equation has $t{-}1$ and earlier valid instruments.

Exogeneity: strict exogeneity of x_{it} means the entire time series is a valid set of instruments in each of the first differenced equations in addition to the response variable $t{-}2$ and earlier. In this case, $E(\varepsilon_{it}|x_{is}) = 0 \; \forall \; i = 1, \dots, N, \forall \; t, s = 1, \dots, T$ together with any other instrument, can enter the instrument matrix, **Z**, in FD, with one column per instrument.

Appendix A4.2: Equality of Results from Helmert Transformation of Raw and Demeaned Data

Arellano and Bover (1995) formerly developed a data transformation approach based on the Helmert technique that does not suffer from the gap problem experienced when using the first-difference method. Love and Zicchino (2006) then developed a panel vector autoregression code for stata (pvar2), see also Ryan Decker's Note on the Helmert's transformation. This Appendix proves the equivalence between the results generated from either raw or demeaned data when using the Forward Orthogonal Deviations Transform.

Definitions: Suppose x_{it}^{H} denotes the Helmert-transformed version of raw data for, say, sector i over time t. Then

$$x_{it}^{H} = \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \frac{1}{T-1}\sum_{n=t+1}^{T} x_{in}\right)$$

where $t \in (1,2, \dots, T)$. Notice that x_{it}^{H} for time t is the average of all future observations from time $t+1$ through T. Observe also that this expression weighs heavily for observations closer to the beginning of the time series.

Now consider x_{it}^{HD} to be a time-demeaned Helmert-transformation so that

$$x_{it}^{HD} = \sqrt{\frac{T-1}{T-t+1}}\left(\ddot{x}_{it} - \frac{1}{T-1}\sum_{n=t+1}^{T} \ddot{x}_{in}\right)$$

where $\ddot{x}_{in} = x_{in} - \bar{x}_i$ and $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$.

Proposition: $x_{it}^{HD} = x_{it}^{H}$

Proof:

$$x_{it}^{HD} = \sqrt{\frac{T-1}{T-t+1}}\left(\ddot{x}_{it} - \frac{1}{T-1}\sum_{n=t+1}^{T} \ddot{x}_{in}\right)$$

$$= \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \bar{x}_i - \frac{1}{T-1}\sum_{n=t+1}^{T} (x_{in} - \bar{x}_i)\right)$$

$$= \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \bar{x}_i - \frac{1}{T-1}\sum_{n=t+1}^{T} x_{in} + \frac{1}{T-1}\sum_{n=t+1}^{T} \bar{x}_i\right)$$

$$= \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \bar{x}_i - \frac{1}{T-1}\sum_{n=t+1}^{T} x_{in} + \frac{1}{T-1}(T-1)\bar{x}_i\right)$$

$$= \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \bar{x}_i - \frac{1}{T-1}\sum_{n=t+1}^{T} x_{in} + \bar{x}_i\right)$$

$$= \sqrt{\frac{T-1}{T-t+1}}\left(x_{it} - \frac{1}{T-1}\sum_{n=t+1}^{T} x_{in}\right)$$

$$= x_{it}^{H}$$

Q.E.D.

Appendix A4.3: GMM Estimation of the Structural Equation of Investment using the Principal Component Analysis (PCA) for Reduction of Instrument Proliferation

	GMM DIFF−PCA		GMM SYS−PCA	
Variable	One−Step	Two−Step	One−Step	Two−Step
$\frac{I_{t-1}}{k_{t-2}}$	−0.014	0.025	0.358	0.358*
	(0.2536)	(0.2817)	(0.1839)	(0.1665)
$\frac{s_t}{k_{t-1}}$	−0.187	−0.244	−0.163	−0.162
	(0.1801)	(0.199)	(0.1692)	(0.1584)
$\frac{s_{t-1}}{k_{t-2}}$	−0.051	−0.059	0.215*	0.225*
	(0.1026)	(0.0845)	(0.0969)	(0.1053)
Emp_t	−0.01	−0.006	−0.128	−0.146
	(0.1411)	(0.1642)	(0.1315)	(0.1419)
Emp_{t-1}	0.16	0.185	0.181	0.199
	(0.1385)	(0.1527)	(0.1305)	(0.1389)
Constant	−	−	−0.221	−0.232
	−	−	(0.2193)	(0.2047)
NT	100	100	171	171
N	43	43	68	68
AR(1)−*p-value*	0.062	0.17	0.024	0.069
AR(2)−*p-value*	0.033	0.13	0.047	0.145
Sargan −*p-value*	0.0499	0.0499	0.0348	0.0348
Hansen −*p-value*	0.9992	0.9992	0.967	0.967
#Z	76	76	78	78
#X	12	12	12	12
Wald χ^2 −Test	77.9	47.85	279.87	166.22
χ^2_p	0	0	0	0
h	3	3	3	3

Legend: Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.
Note: The high Hansen p-value suggests that high instrument proliferation caused over-fitting of endogenous variables, see Roodman (2009a, p. 98). The covariate estimates can therefore serve as upper bounds.

REFERENCES

Abel A.B. (2014), ―Analytics of Investment, q and Cash Flow‖, *National Bureau of Economic Research*.

Abel A.B. and J.C. Eberly (1994), ―A Unified Model of Investment under Uncertainty‖, *The American Economic Review*, 84(5): pp. 1369-1384.

Aitkin, M. and M. Alfo. (2003), ―Longitudinal analysis of repeated binary data using autoregressive and random effect modelling‖, *Statistical Modelling*, **3**: pp. 291–303.

Akay A. (2012), ―Finite Sample Comparison of Alternative Methods for Estimating Dynamic Panel Data‖ Models, *Journal of Applied Econometrics* 27: 1189-1204.

Albarran P., R. Carrasco and J.M. Carro (2015), ―Estimation of Dynamic Nonlinear Random Effects Models with Unbalanced Panels‖, Mimeo.

Alonso-Borrego C. and M. Arellano (1999), ―Symmetrically Normalized Instrumental-Variable Estimation Using Panel Data‖, Journal of Business and Economic Statistics, 17(1): pp.36-49.

Altonji J.G. and L.M. Segal (1996), ―Small-Sample Bias in GMM Estimation of Covariance Structures‖, *Journal of Business & Economic Statistics*, 14(3): pp. 353-366.

Anderson T. W. And C. Hsiao (1981), "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, 76: pp.598-606.

Anderson, T. W. And C.Hsiao (1982), "Formulation and Estimation of Dynamic Models Using Panel Data", *Journal of Econometrics*, 18: pp.47-82.

Arellano M. and S. R. Bond (1991), ―Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations‖, *The Review of Economic Studies*, 58(2): pp. 277-297

Arellano, M., and O. Bover (1995), ―Another Look at the Instrumental Variable Estimation of Error-Components Models‖, *Journal of Econometrics*, 68: pp. 29-51.

Bai J. and S. Ng (2010), ―Instrumental Variable Estimation in A Data Rich Environment‖, *Econometric Theory*, 26: pp. 1577–1606.

Bigsten A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, R. Oostendorp, C. Pattillo, M. Soderbom, and F. Teal (2005), ―Adjustment Costs and Irreversibility as Determinants of Investment: Evidence from African Manufacturing‖, *Contributions to Economic Analysis & Policy*,4(1).

Bloom N. (2009), ―The Impact of Uncertainty Shocks‖, *Econometrica*, 77(3): pp.623–685.

Blundell and Bond (1998), ―Initial Conditions and Moment Restrictions in Dynamic Panel Data Models‖, Journal of Econometrics (87): pp. 115-143.

Blundell R., S. R. Bond and F. Windmeijer (2000), ―Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator‖, Institute for Fiscal Studies, WP 00/12.

Blundell R.W. and S.R.Bond (2000), ―GMM Estimation with Persistent Panel Data: An Application to Production Functions‖, *Econometric Reviews* (19): 321-340.

Blundell, R., and S. Bond (1998), ―Initial Conditions and Moment Restrictions in Dynamic Panel Data Models‖, *Journal of Econometrics*, 87: pp. 115-143.

Bock, R.D. and Lieberman, M. (1970), ―Fitting a response model for n dichotomously scored items‖. *Psychometrika*, 33: 179–197.

Bond S.R. (2002), ―Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice‖, Institute for Fiscal Studies, WP.CWP09/02.

Bontempi M. E. and R. Golinelli (2014), ―Panel Data Econometrics: Theory and Applications in Stata Dynamic Models for Panel Data‖, Lecture Notes.

Bontempi M.E. (2015), ―Investment–Uncertainty Relationship: Differences between Intangible and Physical Capital‖, *Economics of Innovation and New Technology*, 2: pp.1-29.

Bontempi M.E. and I. Mammi (2015), ―Implementing a Strategy to Reduce the Instrument Count in Panel GMM‖, *The Stata Journal*, 15(4): 1075-1097

Bowsher C. G. (2002), ―On Testing Overidentifying Restrictions in Dynamic Panel Data Models‖, Mimeo.

Bun M.J.G and F. Windmeijer (2010), ―The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models‖, *Econometrics Journal*. 13: 95-125

Caballero R. J. and E.M.R.A. Engel (1999), ―Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach‖, Econometrica, 67(4): pp.783-826.
Caballero R.J. and J.V. Leahy (1996), ―Fixed Costs: The Demise of Marginal q‖, National Bureau of Economic Research, WP.5508.
Caballero R.J., E.M.R.A. Engel and J.C. Haltiwanger (1995), ― Plant-Level Adjustment and Aggregate Investment Dynamics‖, *Brookings Papers on Economic Activity*, (2): pp.1–39.
Cameron A.C. and P.K. Trivedi (2005), ―Microeconometrics: *Methods and Applications*‖, Cambridge University Press, Cambridge.
Cameron A.C. and P.K. Trivedi (2009), ―Microeconometrics using Stata‖, College Station, TX: Stata Press.
Chamberlain,G. (1984), ―Panel data‖, In *Handbook of Econometrics*, vol. II (eds Z. Griliches and M.D. Intriligator), pp. 1247–1318. Amsterdam: North-Holland.
Chay K.Y. and D.R. Hyslop (2000), ―Identification and Estimation of Dynamic Binary Response Panel Data Models: Empirical Evidence using Alternative Approaches‖,
Christiano L., M. Eichenbaum and C. Evans (2005), ―Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,‖ *Journal of Political Economy*, 113: 1—45, 2005.
Chrysanthou G.M. (2008), Determinants of Trade Union Membership in Great Britain During 1991-2003
Cooper R. W, J.C. Haltiwanger and L. Power, (1999), ―Machine Replacement and the Business Cycle: Lumps and Bumps‖, *The American Economic Review*, 89(4): 921-946.
Cooper R.W. and J. C. Haltiwanger (2006), ―On the Nature of Capital Adjustment Costs‖, The Review of Economic Studies, Vol. 73, No. 3 (Jul., 2006), pp. 611-633.
Drakos K. and P. Konstantinou (2013), ―Investment Decisions in Manufacturing: Assessing the Effects of Real Oil Prices and their Uncertainty‖, Journal of Applied Econometrics, 28(1): pp.151-163.
Dutoit L.C. (2007), ―Heckman's Selection Model, Endogenous and Exogenous Switching Models: A survey‖, Mimeo.
Eberly J., S. Rebelo and N. Vincent (2012), ―What Explains the Lagged Investment Effect?‖ Mimeo.
Erikson T. and T.M. Whited (2000), ―Measurement Error and the Relationship between Investment and *q*‖, *Journal of Political Economics*, 108(5): pp.1027-1057.
Gala V. (2015), ―Measuring Marginal q‖, Mimeo.
Galle S. (2016), ―Competition, Financial Constraints and Misallocation: Plant-Level Evidence from Indian Manufacturing‖, Mimeo, University of California, Berkeley.
Gelos R.G. and A. Isgut (2001), ―Fixed Capital Adjustment: Is Latin America Different?‖, *The Review of Economics and Statistics*, 83(4): pp. 717-726.
Gilchrist and Himmelberg (1998), ―Investment: Fundamentals and Finance‖, *NBER Macroeconomics Annual*, Vol. 13 (1998), pp. 223-262.
Goldfeld S.M. and R.E. Quandt (1973), ―Estimation of Structural Shift by Switching Regression‖, *Annals of Economic and Social Measurement*, 2: pp.475-485.
Gunning J.W. and T. Mengistae (2001), "Determinants of African Manufacturing Investment: The Microeconomic Evidence", *Journal of African Economies*, AERC Supplement 2, pp. 48-80.
Han C. and P.C.B. Philips (2010), ―GMM Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity‖, *Cowles Foundation for Research in Economics*.
Hansen L.P. (1982), ―Large Sample Properties of Generalized Method of Moments Estimators‖, Econometric, 50(4): pp.1029-1054.
Hayakawa K. (2009), ―A Simple Efficient Instrumental Variable Estimator for Panel AR(p) Models When Both N and T are Large‖, *Econometric Theory*, 25(3): pp. 873-890.
Hayakawa K. (2014), Asymptotic Properties of the System GMM Estimator in Dynamic Panel Data Models When Both N and T are Large‖, Mimeo, Hiroshima University.
Heckman and Singer (1982), ―The Identification Problem in Econometric Models for Duration Data.‖ invited paper in *Advances in Econometrics, Proceedings of the Fourth World Congress of the Econometric Society*, ed. by W. Hildenbrand. Cambridge: Cambridge University Press, pp. 39-77.
Heckman J.J. (1981a), ―The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process‖, In *Structural Analysis of Discrete Data*

with Econometric Applications (eds C. F. Manski and D. L. McFadden), pp. 179–195. Cambridge: MIT Press.
Heckman J.J. (1981b), ―Heterogeneity and state dependence‖, In *Studies in Labour Markets* (ed. S. Rosen), pp. 91–139. Chicago: University of Chicago Press.
Heckman J.J. and B. Singer (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52 (2): 271-320.
Hu X. and F. Schiantarelli (1998), ―Investment and Capital Market Imperfections: A Switching Regression Approach Using U.S. Firm Panel Data‖, *The Review of Economics and Statistics*, 80(3): 466-479.
Huh K. and R.C. Sickles (1994), ―Estimation of the Duration Model by Nonparametric Maximum Likelihood, Maximum Penalized Likelihood and Probability Simulators‖, *The Review of Economics and Statistics*, 76(4): pp. 683-694.
Hyslop, D. R. (1999), ―State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labour Force Participation Behaviour of Married Women,‖ *Econometrica*, 67, 1255-1294.
Judson R. A. and A.L. Owen (1999), ―Estimating dynamic panel data models: a guide for macroeconomists‖, *Economics Letters*, 65: pp. 9–15.
Kapetanios G. and M. Marcellino (2007), ―Factor-GMM Estimation with Large Sets of Possibly Weak Instruments‖, Mimeo.
Kiviet J.F., M. Pleus and R.W. Poldermans (2017), "Accuracy and Efficiency of Various GMM Inference Techniques in Dynamic Micro Panel Data models", *Econometrics*, Vol.5 (14), pp. 1-54.
Lancaster T. (1979), "Econometric Methods for the Duration of Unemployment," *Econometrica* 47: pp.939-956.
Lancaster T. and S. Nickell (1980), "The Analysis of Reemployment Probabilities for the Unemployed," *Journal of the Royal Statistical Society*, 143: pp.141-165.
Lee L. F. and R.H. Porter (1984), ―Switching Regression Models with Imperfect Sample Separation Information With an Application on Cartel Stability‖, *Econometrica*, Vol. 52, No. 2: pp. 391-418.
Lee L.F. and R.P.Frost (1978), ―Estimation of Some Limited Dependent Variable Models with Applications to Housing Demand‖, Journal of Econometrics, 8: pp.357-382.
Letterie W.A. And Pfann (2007), ―Structural Identification of High and Low Investment Regimes‖, *Journal of Monetary Economics*, 54: pp. 797–819.
Lindsay, B. G., (1983), "The Geometry of Mixture Likelihoods: A General Theory," *The Annals of Statistics*, 11: pp.86-94.
Lokshin M. and Z. Sajaia (2004), ―Maximum likelihood estimation of endogenous switching regression models‖, *The Stata Journal*, 4(3): 282–289.
Love I. and L. Zicchino (2006), ―Financial development and dynamic investment behaviour: Evidence from panel VAR‖, The Quarterly Review of Economics and Finance, Vol.46: pp.190–210.
Maddala, G.S. (1983), ―Limited Dependant and Qualitative Variables in Econometrics‖, New York: Cambridge University Press.
McClelland J. (1997), ―"Permit Renewal and the Timing of Major Investment", Mimeo, University of Maryland.
Mehrhoff, J., (2009) ―A solution to the problem of too many instruments in dynamic panel data GMM‖, *Discussion paper n. 1/2009*, Deutsche Bundesbank.
Mundlak, Y. (1978), ―On the pooling of time series and cross section data‖, *Econometrica*, 84: pp. 69–85.
Nabi I. (1989), ―Investment in Segmented Capital Markets‖, *The Quarterly Journal of Economics*, 104(3): pp. 453-462.
Nilsen O. A., and F. Schiantarelli (2003), ―Zeros and Lumps in Investment: Empirical Evidence on Irreversibilities and Nonconvexities‖, *The Review of Economics and Statistics*, 85(4): pp.1021-1037.
Rabe-Hesketh S. and A. Skrondal (2012), ―*Multilevel and Longitudinal Modelling Using Stata*‖, StataCorp LP, College Station, Texas.
Rabe-Hesketh S. and A. Skrondal (2013), ―Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem‖, *Economic Letters*, 120**:** pp.346–349.

Rabe-Hesketh, S., A. Pickles, and A. Skrondal, (2003), ―Correcting for Covariate Measurement Error in Logistic Regression using Nonparametric Maximum Likelihood Estimation‖, *Statistical Modelling,* **3:** pp.215–232.
Roodman D. (2009a), ―How to Do Extabond2: An Introduction to difference and system GMM in Stata‖, *The Stata Journal,* 9 (4): pp.86-136.
Roodman D. (2009b), ―Practitioners' Corner: A Note on the Theme of Too Many Instruments‖, *Oxford Bulletin of Economics and Statistics,* 71: pp.135-158.
Seaman *et al* (2013), ―What is meant by ―Missing at Random‖?‖ *Statistical Science*, 28(2): 257–268.
Skrondal A. and S. Rabe-Hesketh (2014), ―Handling Initial Conditions and Endogenous Covariates in Dynamic/Transition Models for Binary Data with Unobserved Heterogeneity‖, *Applied Statistics* 63(2): pp.211-237.
Stewart M. (2006), ―Maximum Simulated Likelihood Estimation of Random-Effects Dynamic Probit Models with Autocorrelated Errors‖, The Stata Journal, 6(2): pp.256-272.
Stewart M. (2007), ―The Interrelated Dynamics of Unemployment and Low-Wage Employment‖, *Journal of Applied Econometrics,* 22: pp.511–531.
Tauchen G. (1986),‖Statistical Properties of Generalized Method of Moments Estimators of Structural Parameters Obtained from Financial Market Data‖, *Journal of Business and Economic Statistics,* 4: pp.397-416.
Windmeijer F. (2005), ―A Finite Sample Correction for the Variance of Linear Two-Step GMM Estimators‖, *Journal of Econometrics,* 126: pp.25-51.
Wooldridge J. M. (2010), ―Correlated Random Effects Models with Unbalanced Panels‖, Mimeo.
Wooldridge J.M. (2005), ―Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity‖, *Journal of Applied Econometrics*, **20:** pp.39–54.
Wooldridge, J. M. (2010a), *Econometric Analysis of Cross Section and Panel Data,* 2nd edition. Cambridge: MIT Press.
Youssef A, H., A.A. El-Sheik and M.R. Abonazel (2014), ―Improving the Efficiency of GMM Estimators for Dynamic Panel Models‖, MPRA.
Youssef A.H. and M.R. Abonazel (2015), ―Alternative GMM Estimators for First-Order Autoregressive Panel Models: An Improving Efficiency Approach‖, MPRA.
Ziliak (1997), ―Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment condition estimators‖, *Journal of Business and Economic Statistics,* 15: pp. 419-431.