Essays on Industrial Dynamics: Evidence from Swaziland

By Samuel Vika Mhlanga

Dissertation presented for the degree of **Doctor of Philosophy** in the Faculty of Economic and Management Sciences at Stellenbosch University.

Supervisor: Professor Neil A. Rankin

December 2017

DECLARATION

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2017

ABSTRACT

This thesis is completely based on a unique and rich establishment-level panel dataset that has never been used before provided by the Central Statistical Office (CSO) of Swaziland to study industrial dynamics. It begins with an assessment of aggregate resource flows among sectors of the Swazi economy to understand the nature of structural change over a period of 10 years since 1994. We find a slight shift in output and labour from the high-productivity manufacturing to low-productivity agriculture and services sectors, potentially developing into what is also known as the manufacturing hollowing out phenomenon. Within the manufacturing sector itself, the evolution of firm-size distribution appears to converge to a bimodal structure; while deeper investigation produces a missing 'missing middle' in the economy. The analysis goes on to evaluate the job creating prowess of small firms. The general finding again is that job destruction dominates job creation, regardless of firm-size category. However, large firms destroy and create more jobs than small firms, even without relevant data to control for firm age. This suggests an absence of transition channels from subsistence to transformational entrepreneurship in the Swazi manufacturing sector.

An in-depth analysis of the drivers of aggregate productivity growth is also carried out. It is found that resource reallocation across firms is productivity enhancing while longitudinal technical efficiency is productivity reducing in the manufacturing sector. However, the firm entry-exit dynamic is the main contributor to aggregate productivity growth. In the case of investment dynamics and unobserved heterogeneity, there is neither significant impact of the lagged investment variable nor presence of individual firm-specific heterogeneity that might raise firms' propensity to invest in plant, machinery and equipment. That is, the impact of unobserved firm-specific characteristics underlying investment decisions is also insignificant. The most interesting finding though is that the cost of uncertainty in a trade liberalization environment can also be measured in our framework. We find that missing investments at time $t-1$ reduce the propensity to invest at time t by a significant margin. Furthermore, by interacting missing investments with labour, we estimate a significant probability of capital substitution for labour. When an endogenous switching regime model of investment is estimated, the *single* investment regime produced by fixed and random effects models is validated.

OPSOMMING

Hierdie tesis is volledig gebaseer op 'n eiesoortige en omvangryke stel paneeldata op die vlak van ekonomiese instellings wat deur die sentrale statistiekkantoor van Swaziland verskaf is vir die bestudering van industriële dinamika, onder meer totale groei in produktiwiteit. Die studie begin met 'n evaluering van die saamgevoegde vloei van hulpbronne tussen sektore van die Swazilandse ekonomie ten einde insig te verwerf in die aard van strukturele veranderings wat oor 'n tydperk van tien jaar sedert 1994 plaasgevind het. Ons sien 'n effense verskuiwing in uitset en arbeid, van die hoëproduktiwiteit-vervaardigingsektor na die laeproduktiwiteitsektore van landbou en dienslewering, om moontlik te ontwikkel in die verskynsel wat as die uitholling van vervaardiging bekend staan. Binne die vervaardigingsektor self ontwikkel die verspreiding van ondernemingsgrootte in die rigting van 'n bimodale struktuur; 'n aanduiding van 'n "vermiste middel" in die ekonomie. Die ontleding gaan voort deur die vermoë van kleinsakeondernemings om werk te skep te ontleed. Die bevinding in die algemeen is weer eens dat werksgeleenthede wat vernietig word die werksgeleenthede wat geskep word oorheers, ongeag die kategorie van ondernemingsgrootte. Groot ondernemings vernietig én skep egter meer werksgeleenthede as kleiner ondernemings, selfs sonder tersaaklike data om vir die ouderdom van ondernemings te kontroleer. Dit dui op die afwesigheid van kanale wat ondernemings in die Swazilandse vervaardigingsektor van bestaansentrepreneurskap na transformasionele entrepreneurskap kan laat oorgaan.

'n Diepgaande ontleding van die drywers van totale produktiwiteitsgroei word ook gedoen. Daar word bevind dat die toedeling van hulpbronne oor ondernemings heen produktiwiteit verbeter, terwyl longitudinale tegniese doeltreffendheid produktiwiteit in die vervaardigingsektor laat afneem. Die toetree-uittree-dinamiek van maatskappye is egter die grootste bydraer tot totale produktiwiteitsgroei. In die geval van investeringsdinamika en onopgemerkte heterogeniteit is daar nie 'n beduidende impak van die vertraagde investeringsveranderlike of die aanwesigheid van individuele, ondernemingspesifieke heterogeniteit wat ondernemings se geneigdheid om in masjinerie en toerusting te investeer, moontlik sal verhoog nie. Dit wil sê, die impak van onopgemerkte ondernemingspesifieke eienskappe onderliggend aan investeringsbesluite is ook onbeduidend. Die interessantste bevinding is dat die koste van onsekerheid in 'n omgewing van handelsliberalisering ook in ons raamwerk gemeet kan word. Ons vind dat verlore investering op tydstip $t-1$ die geneigdheid om te investeer by tydstip t met 'n aansienlike marge laat afneem. Deur verlore investering voorts met arbeid in wisselwerking te plaas, dui ons beraming op 'n beduidende waarskynlikheid dat arbeid met kapitaal vervang word. 'n Beraamde model van 'n endogene omruilingsbestel bekragtig die *enkelvoudige* investeringsbestel wat deur modelle van vaste en ewekansige effekte geproduseer word.

<u>EXAMINATION COMMITTEE</u>

Professor Willem Boshoff -   Chairman, University of Stellenbosch

Professor Rulof Burger   -   University of Stellenbosch

Professor James Fairburn -   University of Kwazulu Natal

Professor Waldo Krugell -   University of the North West

Professor Neil A. Rankin -   Thesis Advisor, University of Stellenbosch

DEDICATION

This work is dedicated to the Word. The One who was, is and is to come. The One who is holy, the One who is true, the One who holds the key of David. It is He that openeth and no man shutteth; and He that shutteth and no man openeth. He is the Alpha and the Omega, the beginning and the end, the first and the last. Glory be to our Lord and Saviour Jesus Christ, the Messiah! *Amen*.

TABLE OF CONTENTS

CHAPTER 3

CHAPTER 4

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

MOTIVATION AND STRUCTURE

The manufacturing sector traditionally plays an important role as economies grow and industrialize, contributing to the overall gross domestic product (GDP), and to the growth of productivity and employment. Its performance and dynamics continue to preoccupy economists, policymakers, and the public. However, economic development through industrialization seems to be becoming more difficult; see Rodrik (2006a, 2013) and McMillan and Rodrik (2011). A deeper understanding of the performance of the industrial sector can help small developing countries like Swaziland design policies which take these constraints to economic development into account.

It is often the case that empirical analysis of the sector is performed at the aggregate level, which masks the heterogeneous behaviour of firms, including the churning that characterizes the labour market and firm turnover. Every year jobs are created while others are destroyed as firms expand and contract, or new firms enter the market while old ones shut down businesses. At high levels of aggregation, differentials in the magnitude of plant-level output growth induced by an additional unit of labour effort are hard to quantify. Aggregates prevent any analysis of the impact that the underlying firm's entry-exit dynamics have on investment, and prohibit an estimation of observed and unobserved micro-effects on investment rates. Understanding micro-aspects matters, since sector and economy-wide outcomes are an aggregation of firm-level activities. Furthermore, appropriate policy responses might differ depending on the nature of firm-level behaviour. That is why theoretical and micro-econometric research has increased since the 1980s in response to improved access to firm-level datasets which help researchers produce sharper results.

As part of their general remit, many government statistical agencies, including the Central Statistical Office (CSO) of Swaziland, collect annual firm-level census data purely for internal office use. This is particularly for the calculation of aggregate measures such as the National Accounts. These datasets, if they are made more broadly available to researchers, can provide the basis for understanding both the firm-level dynamics and macroeconomic outcomes. This thesis represents the first effort to use micro-data collected by the Swaziland CSO to investigate industrial dynamics during a key period: the democratic transition and subsequent liberalisation in South Africa, a country to which Swaziland is inextricably linked.

Benchmarked against the U.S. Longitudinal Research Database (LRD), the Swazi dataset is of good quality in terms of coverage and measurement of variables of interest. Although deficient in certain respects, it allows for the analyses of job movements in the labour market, firm turnover, investment dynamics and input productivity growth. An interesting aspect of this work is that it is on a small landlocked economy surrounded by a trading partner in the Southern African Customs Union (hereinafter referred to as the Customs Union) that is geographically 70.26 times larger and in 1994 was economically 53.42 times larger. This means that in many ways Swaziland is not an equal partner

in the Customs Union and has to abide by economic decisions taken by larger members (i.e. South Africa). A recent period of substantial change occurred when South Africa democratized in the early 1990s. This democratization was accompanied by substantial changes in the trade regime which affected the Customs Union. The dataset itself spans the entire period of trade liberalization precipitated by South Africa's reintegration into the world economy in the mid-1990s and thus can be used to investigate how Swazi firms responded to this course of events.

This new economic environment exposed producers in the Customs Union to more import competition in a similar fashion to what occurred under similar circumstances in countries like Chile (see Pavcnik, 2002). As demonstrated by Edwards and Behar (2005), the exposure that establishments had to import competition in South Africa increased their access to new foreign technology that enhanced the innovation aspect of productivity in domestic industries. Trade liberalization led to the loss of domestic producer market share in the region and to expansion of output induced by exploitation of scale economies, particularly in the larger trading partner's market.

In the context of the Swazi economy, however, gains from economies of scale are improbable since increasing returns might be associated with industries involved in import competition. If trade reforms reduce market share of domestic firms without an increase in foreign exports, their propensity to invest in foreign technology is likely to decline as protection comes to an end. Therefore, the benefits of cheaper capital import goods and access to foreign technology made available by tariff reductions are eroded. Although these economic reforms aid procurement of foreign technology, it is uncertain if domestic firms adopt such innovations. Some models show how the benefits of innovation are spread from one country to another either through knowledge transfer or through the exchange of goods. The compelling finding in this case is that the effect of technology diffusion on productivity is vitally dependent on the proximity of the technology source and how flexible the labour market is.

Furthermore, firm-level heterogeneity in its different dimensions suggests that trade liberalization may enhance the productivity of firms by inducing primary input and output reshuffling from inefficient to more productive firms within the same industry. Firm dynamics such as business shutdowns may contribute significantly to the resource reallocation process. In particular, high tariff barriers permit the coexistence of establishments with varying levels of productivity. Dismantling these trade barriers lowers domestic prices, thereby driving high-cost manufacturers out of business. However, these productivity gains are available only if the disposal of capital investment is easy enough not to hinder the exit process of less productive plants.

In some instances, low productivity firms may opt not to exit the market but rather to engage in business reconfiguration in order to improve productivity and confront the new competition brought about by trade tariff reductions. Even if trade liberalization enhances productivity through the various channels, it may achieve that at the cost of firm exit, large resource reallocations and displacement of

primary inputs. The threat of initial costs of worker displacements and business closures deters policymakers from opening up their domestic markets to foreign competition. In cases where the option to choose whether or not to liberalize trade is unavailable, one likely policy choice involves pursuing programmes that promote firm entry. An example of this is the Swazi factory shell construction programme that reduces the fixed costs incurred by new manufacturers; see Ministry of Economic Planning and Development Report (2004/2005).

This thesis contains four chapters dedicated to understanding the manufacturing sector in Swaziland, and is a unique contribution given the nature of the dataset and the time period it covers. This contribution is both Swaziland-specific, and also adds to the broader literature on firm-level dynamics. Although the thread of interconnection between consecutive chapters is embedded in our approach, each chapter is presented as a paper suitable for journal publication. As a prelude to the study of behaviour of primary inputs in manufacturing, we first consider the aggregate inter-sectoral movement of resources to determine the pattern of structural change in Chapter 1. In constructing a transition from the analysis of macroeconomic variables to the analysis of behavioural patterns of plant-level resources, we clean up the data, define variables and evaluate the panel dataset for quality assurance.

In Chapter 2, we investigate the patterns of job flows to determine the role of small firms in creating jobs in the manufacturing sector. The chapter begins with descriptive analyses of employment trends for each two-digit ISIC industry and studies the evolution of firm-size distribution to extract some stylized facts about the sector. After laying out the precise framework for in-depth analysis, it goes straight to the empirical analysis of job flows. We then conclude with linking job flows to industrial productivity in order to learn about the role of turnover on aggregate labour productivity (ALP) growth.

Chapter 3 is concerned with measurement issues associated with the decomposition and analysis of ALP growth. It outlines conventional methods of estimating the drivers of ALP growth and contrasts them with a new approach based on micro-foundations. The latter approach tracks the value of the marginal product (VMP) of labour and uses the index number theory to estimate the right-hand side components of aggregate productivity growth (APG). The ultimate goal in this analysis is to establish whether or not effects of technology diffusion dominate the impact of input reallocation across firms in manufacturing. To the best of our knowledge, this is the first study to use the VMP to estimate APG in an African economy.

Finally, Chapter 4 estimates a structural model of investment to determine the potential role of state dependence and the impact of unobserved heterogeneity. It also investigates whether or not firms self-select into high or low investment regimes. All these objectives are achieved by using fixed-effects methods, an array of random-effects techniques as well as regime switching regressions to produce

the answers. This Chapter uses a novel approach based on hierarchical modelling techniques to piece together random-effects models and nonparametric maximum likelihood estimation methods (NPMLE). Such an approach is used for the first time in a structural model of investment that also provides a framework to estimate the cost of delayed investment.[1]

The slow pace of firm-level investment in capital goods appeared pronounced in the data, implying a potentially high level of economic uncertainty during the period under study. This suggests a need for firms to exercise an option-to-wait strategy. An economic environment characterized by high levels of uncertainty at time t reduces the probability of investing in PME at time $t + 1$. This cost of the option-to-wait strategy can be too high for the sector to experience optimal growth. The strategy also implies some capital/labour substitution measured by the sensitivity of investment rates to changes in employment.

As a whole, the thesis suggests that the Swazi economy experienced a lacklustre performance and a potential hollowing-out process in the manufacturing sector during the period 1994-2003. There is evidence of resources reallocating from high-productivity manufacturing to low-productivity agriculture and services sectors. At the plant-level, the manufacturing sector appears to have evolved from a unimodal to a bimodal firm-size distribution by 2003, suggesting a 'missing middle' problem. Firms also destroyed more jobs than they created. Contrary to popular belief, the job creating ability of small industrial firms simply failed. Furthermore, in a decomposition of industrial aggregate productivity growth, technical efficiency effects only worked to reduce productivity growth, while input reallocation across plants was growth enhancing. At the same time, the analysis of investment patterns over the same period showed a high level of inactivity and, contrary to conventional wisdom, the lagged investment variable had no influence on the current rate of investment. Firms themselves were indistinguishable in terms of unobserved investment behaviour; that is, there is no impact of unobserved heterogeneity on investment decisions. However, the measured cost of investment uncertainty was very high, and so was the capital/labour substitution in favour of the latter.

---

[1] I am grateful to Jesús Carro of the Universidad Carlos III de Madrid for his suggestion that I also consider an alternative estimation method based on Minimum Distance techniques as a robustness check.

CHAPTER 1: Overview

1.1   Introduction

Swaziland is a small landlocked and open economy surrounded largely by South Africa and, to a lesser extent, by the Republic of Mozambique. Since 1910, it has been a member of a constellation of five Sub-Saharan countries that form the Southern African Customs Union (SACU); namely, Botswana, Lesotho, Namibia, Swaziland and South Africa. It has also been a member of the Common Monetary Area (CMA) since 1974 involving the same countries, but Botswana. This arrangement grants Swazi exports free market access within the SACU and the CMA sub-regions without incurring any cost of currency exchange. At the same time, some of the country's key commodities enjoy free access to more distant foreign markets such the European Union under the African, Caribbean and Pacific (ACP)-European Commission (EC) cooperation agreement and to the US under the African Growth and Opportunity Act (AGOA) of 2001. These export products are viewed as politically sensitive since they are major drivers of gross domestic product (GDP) through industrial policy that promotes job creation and investment in the manufacturing sector.

Such patterns of development are typical in developing economies. Industrialization in particular has been characterized by positive growth driven largely by structural change since the mid-1980s in many African economies; see McMillan, Rodrik and Verduzco-Gallo (2014), Rodrik (2014), and Timmer, de Vries and de Vries (2014). The structural change component of aggregate productivity growth entails reallocation of input resources across sectors, as opposed to the other component that involves growth induced by within-firm technical change. A few leading development economists such as Young (2012) and Rodrik (2014) have described this period as an 'African Growth Miracle'. It replaces the traditional pessimism of growth prospects with stories of expanded Chinese investment and positive commodity price movements. However, over-dependence on the external environment, low levels of productivity and constrained private sector investment in globally competitive industries might re-ignite pessimism about the potential to create a sustainable and robust growth path for the economies of Africa (Rodrik, 2014).

This thesis is based on a unique establishment-level panel dataset covering a period of 10 years since 1994 to study industrial job flows, productivity and investment dynamics. The data have never been used before for the analysis of industrial dynamics in Swaziland. This confidential information was provided by the Central Statistical Office of Swaziland, or the CSO. Although the source records were largely available electronically, some statistics were in physical form and needed digitization. The Private Enterprise Development in Low-Income countries (PEDL), a joint research initiative of the Centre for Economic Policy Research (CEPR) and the Department for International Development (DFID), assisted with funding for the digitization of the data, hiring of a research assistant, financing of buy-out time at the University of Swaziland and travel costs for research purposes.

The purpose of this investigation is broadly two-fold. *First*, it exploits the unique dataset to extract evidence on micro-activities that culminate in macro-outcomes during a very interesting period in the Customs Union. *Second*, it answers very specific questions:

a) What is the general nature of structural change in the Swazi economy? How has firm-size distribution in the manufacturing sector evolved? Is the popular belief about the job creating prowess of small firms a valid proposition for the manufacturing sector in Swaziland?

b) What impact does firm-level technical efficiency and primary input reallocation across firms have on aggregate productivity growth (APG) in manufacturing? As an auxiliary question, how much impact does firm turnover have on APG in the sector?

c) What are the characteristic patterns of industrial investment in plant, machinery and equipment in Swaziland? What effects do state dependence and unobserved heterogeneity have on investment decisions? Is a structural investment model best explained in terms of an investment regime switching model in the manufacturing sector? How can the cost of exercising the investment option to wait be measured in an economic environment replete with uncertainty?

A robust finding in the large and growing literature using labour force surveys and population censuses is that trade liberalization has facilitated labour reallocation from inefficient uses to more productive sectors. In Sub-Saharan economies, however; globalization seems to have generated results that move resources from highly productive to less productivity activities, see McMillan and Harttgen (2014) and Rodrik (2013, 2015). This suggests that labour resources are moving from urban factories to country-side agricultural activities or even to informality. Such forms of structural change engender a process of hollowing out, although conceptual and measurement issues around that are not yet settled; see Levinson (2016). At the same time, these countries are said to be characterized by a 'missing middle' where firm-size distributions are bimodal rather unimodal, see Gelb, Meyer and Ramachandran (2014) and Mazumdar and Sarkar (2008).[2]

Similarly, conventional wisdom since the 1980s claims that small firms are principal creators of jobs in market economies and developing nations alike following the empirical work of Birch (1987). Policymakers have responded by designing policies to prop up small firm participation in the economy in the hope of getting more jobs created. Subsequent case studies confirmed the Birch findings; see Davis, Haltiwanger and Schuh (1996) for an elaborate discussion. However, recent work led by Davis *et al*. (1996) and Haltiwanger, Jarmin and Miranda (2013) demonstrates that the standard results about the ability of small firms to create disproportionately more jobs than their larger counterparts are based on flawed conceptual and measurement issues. These researchers found that it is firm birth and young firms, that happen to be small, that actually create jobs more than larger ones.

---

[2] See counter arguments from Hsieh and Olken (2014) for the case of the Indian manufacturing sector.

It is also the new and young firms that are subject to high levels of churning relative to old incumbent firms.

This implies a particular role for the entry-exit dynamics, the productivity of incumbent firms and the primary input reallocation across firms on aggregate productivity growth. The literature is inundated with methods largely based on the neoclassical Solow (1957) model that seeks to estimate the drivers of productivity growth. Most notable among these is Baily, Hulten and Campbell (1992) and its derivatives such as Foster, Haltiwanger and Krizan (2001). However, Petrin and Levinsohn (2012) identifies critical deficiencies associated with the neoclassical approaches to estimating the impact of resource reallocation across producers. Instead, their paper rationalizes a proposition that is based on micro-foundations that trace the value of the marginal product of labour. The latter approach has been applied by, among others, Nishida, Petrin and Polanec (2014) to Chile, Slovenia and Colombia.

The fundamentals of economic development also find partial expression in the robustness of industrial investment. Hence, the behaviour of investment in plant, machinery and equipment is also an interesting aspect of this work. In particular, we ask; what is the nature of longitudinal dependence of investment due to the effects to its previous state and dependence due to firm-specific characteristics such as managerial efficiencies; that is, unobserved heterogeneity? A large literature estimating structural models relies on either fixed or random effects with balanced datasets; see Arellano and Bond (1991), Stewart (2007) and Drakos and Konstantou (2013). However, the insistence on estimating dynamic nonlinear models under conditions of balanced datasets leads to the loss of useful information and to estimation difficulties due to potentially insufficient observations. As Albarran, Carrasco and Carro (2015) argue, the problem is magnified in structural investment models with a high incidence of missing values.

This thesis makes four novel contributions to the literature. *First*, it presents the first systematic results on the creation and destruction of jobs in the Swazi manufacturing sector. *Second*, it uses standard approaches and new methods based on micro-foundations to estimate aggregate productivity growth over time and across broadly defined industries. The value of the marginal product has never been used in any African economy before, let alone in a small open economy within a liberalizing customs union. *Third*, it demonstrates the impact of confounding effects of plant turnover on resource reallocation effects estimates calculated on the basis of neoclassical approaches that populate the literature. *Fourth*, it uses firm-level data to estimate a structural nonlinear investment model without the requirement for a balanced dataset. Instead, it relies on a method that does not discard observations and still generate unbiased results concerning the variables of interest. To the best of our knowledge, this estimation method has never been applied before at this level of disaggregation to study the industrial behaviour of producers.

It is therefore instructive to begin this study by providing a brief economic background on the country in question. Swaziland is a small and landlocked open economy surrounded largely by South Africa and, to a lesser extent, by the Republic of Mozambique. Despite its foreign trade index of 1.67 in 2000 and a relatively diversified production structure, the country's economic development appears to be caught in a middle-income trap; see Brixiová and Kangoye (2013) and Edwards *et al*. (2013). Its economic growth has at best stagnated since the 1990s following the lifting of economic sanctions on South Africa and the ensuing *de facto* trade liberalization in the Customs Union.

These economic reforms facilitated industrial structural adjustments where new industries were created and others were destroyed through the firm-exit dynamic. The impact of the observed firm churning and behaviour of incumbent firms on capital goods investment was characterized by caution concerning input procurement. Rather, rational firms raised employment only marginally to keep operations running since the hiring and firing costs are not as costly a proposition as the cost of investment irreversibility.[3] Investment in specific skills required by the manufacturing sector to remain competitive in the Customs Union and beyond was also held back. This put the country at the risk of lagging behind its comparator economies and drifting away from its own development path (Edwards *et al*., 2013). Its exposure to foreign trade shocks due to global and regional economic crises dampened the demand for industrial goods manufactured in Swaziland and therefore produced low economic growth in the sector.

This vulnerability to external events reinforced some of the already identified structural constraints to growth and competitiveness. There is notable heterogeneity in industrial exposure to exogenous shocks and channels for their propagation throughout the sector and the economy as a whole. Following the logic of Gabaix (2011), any negative shocks hitting, for example, the sugar industry which has firms on the right (fat) end of firm-size distribution is likely to affect macroeconomic outcomes. Similarly, given the limited domestic population of firms in each industry, any strategic move between two large firms to either merge or acquire another, shakes up the structure of capital investments and employment of the whole industry.

This portrait of industrial economic growth and development in Swaziland is also documented in the country's macroeconomic performance indicators. Any robust and reliable micro-analysis should therefore be amenable to a form of aggregation that matches these macroeconomic outcomes. In particular, it should mimic the sectoral outputs, fixed capital stock and employment levels. In the next section, the structure of the Swazi economy is assessed. Section 1.3 explains the procedure used in the preparation and management of data. Section 1.4 performs an elaborate demonstration to show that the annual census data on manufacturers is of good quality and therefore suitable for deeper analysis in this thesis.

---

[3] See similar arguments in Bentolila and Bertola (1990).

1.2   The Structure of the Swazi Economy

The concepts of structure and structural change in economics originate from at least two main sources. One such source is based on the dual-economy approach in development economics and was first introduced by Lewis (1954) and Ranis and Fei (1961). It views the economy as a structurally heterogeneous system represented by two sectors: one traditional and the other modern, each with very specific characteristics. In particular, as argued by Rodrik (2013), the traditional sector depends on technologically backward methods of production while the modern one accumulates human and physical capital, innovates, and raises its productivity growth. Economic growth is in this sense essentially an outcome of resource flows from the traditional to the modern sector. The other source has its foundations in macroeconomics under the neoclassical framework of the Solow (1957) growth model. In contrast, this model assumes a constellation of heterogeneous activities which are structurally similar enough to be aggregated in a representative sector, see Rodrik (2013). A typical condition presented in either framework is the assumption of full employment, see Rodrik (2006b).

It seems useful to think of these conceptual insights as complementary perspectives on economic growth, while considering the agricultural and manufacturing/services activities as traditional and modern sectors, respectively. According to Rodrik (2013), this provides a basis for associating the dual-economy principle with inter-sectoral economic relationships and flows which allow skilled labour to move from unsophisticated agriculture to the modern manufacturing/services sector. It also raises two issues. First, structural transformation is derived from the rapid inter-sectoral reallocation of resources. The sophisticated sector is expected to operate under conditions of increasing returns to scale, see Nassif, Feijó and Araújo (2014). The second is the fundamentals challenge of increasing skilled labour and effective institutions needed to support productivity across industries in both manufacturing and services sectors, see Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012) for an argument on the robust impact of institutions on long-run development.

In the African context, data collected by the Groningen Growth and Development Centre as well as the World Bank's World Development Indicators suggest that the agricultural sector has lost labour inputs and value added largely to the services sector rather than to manufacturing since the 1960s (Rodrik, 2014). Specifically, industrialization appears to have lost its vitality since the 1970s without much recovery in the subsequent decades. The countries studied are not sufficiently rich to experience any form of de-industrialization, yet this pattern seems evident in Africa; see McMillan and Rodrik (2011) and Rodrik (2014).

Since the 1990s, developing countries have generally become more integrated into the world economy. A country's ability to benefit from globalization effects appears mostly dependent on its readiness to internalize the technological transfers and associated production efficiencies; see McMillan *et al*. (2014). In readier countries, high productivity jobs have increased and structural

change has contributed to economic growth. In Africa, contrary to conventional wisdom, labour resources pre-2000s have moved from high productivity to low productivity areas, and also to informality (McMillan *et al*., 2014). This was an atypical response of these economies to the standard productivity-enhancing effects of trade liberalization. It was characterized by import competing industries losing low-productivity firms through exit and gaining high-productivity ones through entry. Tariff reduction also required firm-level rationalization of resources by shedding labour to improve production efficiency. However, it is these newly unemployed workers that moved to agriculture and informality.

In the post-2000 period, the economic performance of the African continent is referred to as 'the African Growth Miracle' (Young, 2012) based on the consumption growth rate ranging from 3.4 to 3.7 percent. McMillan *et al*. (2014) found a turning point in the structural change performance of African countries. A positive contribution of labour reallocation from inefficient activities to more productive activities was a prominent characteristic of their results. According to McMillan and Harttgen (2014), the patterns of structural change in Sub-Saharan Africa post-2000 mimic patterns of structural change that characterize the situation in well-functioning market economies. This component of productivity growth contributed one percentage point to aggregate labour productivity in this region.

As a country in the Sub-Saharan region, Swaziland's economic structure and its evolution can be viewed through the lens of a dual economy that is primarily subject to external influences. Since the period of analysis covers the entire trade liberalization episode of 1994-2004, the expectation is a substantial movement of primary inputs and market shares from low-productivity to high-productivity sectors. In Figure 1.1(a), sectoral shares of GDP at factor cost are presented.[4] The agricultural sector experienced a roughly stagnant share of GDP, with a moderate and intermittent annual increase and decline. Although the increase in output shares was evident in both services and manufacturing sectors, the two economies co-moved in output share growth until 1997. After this period, the sectors formed a bell-shaped funnel, with conservative growth in the annual shares of the services' GDP. Specifically, the output share of the manufacturing sector stagnated at 41 percent in the first half of the period and began a steady decline down to 39 percent GDP share by 2003. In summary, as shown in Appendix A1.1, the agricultural share of output is largely fixed at the same level throughout the period of analysis, the manufacturing sector's share of GDP is trending downwards and the services sector's share of GDP is trending upwards. The pattern of economic development in Swaziland mimics global patterns of structural change as shown in Figure 1.1(b). The world industrial and agricultural sectors started growing more slowly than the services sector since 1980.

---

[4] All output time series in Swaziland should be interpreted with caution, given the recent rebasing exercise by the Central Statistical Office from 1985 to 2000 constant prices, which shows a dramatic economic growth rate of 35.5 percent.

Figure 1.1(a): Sectoral Composition of Swazi GDP at 1985 Factor Cost in Emalangeni (LCU)[5]
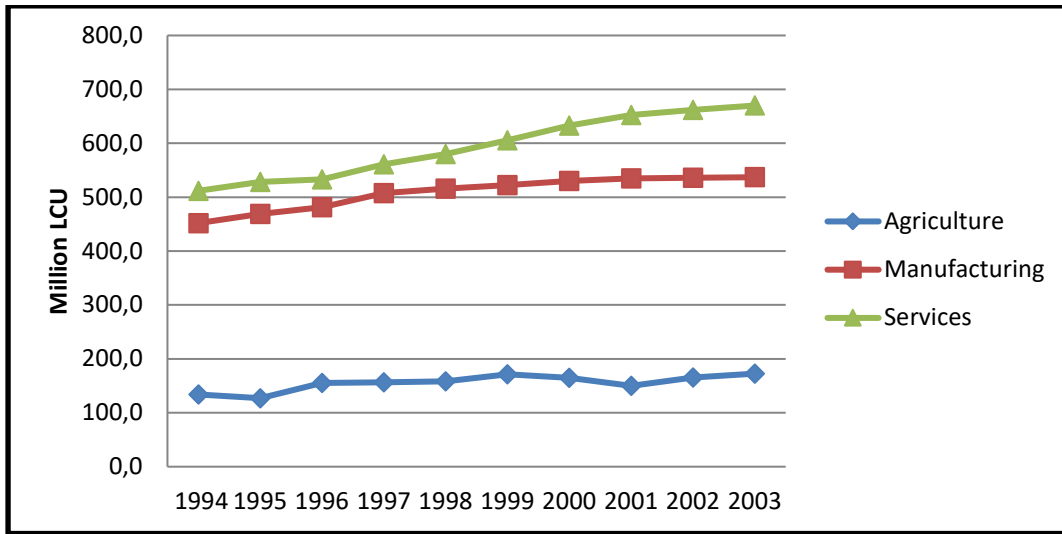
Source: IMF Country Reports 99/13, 00/113 and 06/109.

Figure 1.1(b): Sectoral Composition of World Value Added
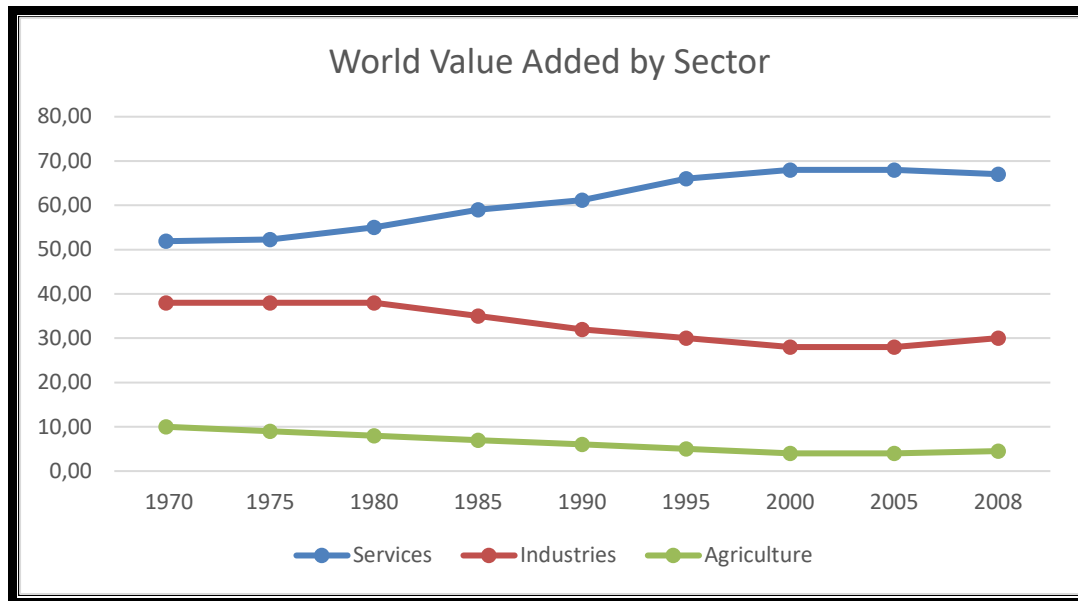
Source: UNIDO calculation based on UN Statistics (data in current prices, in US$).

This empirical outcome is a typical indication of limited structural change in an economic environment that lacks robust industrialization. The size of the manufacturing sector and the degree of global competitiveness of industrial investment may partly explain poor performance in the sector, see Rodrik (2014). This is consistent with the observation by Rodrik (2014) that economic development in Africa is not likely to come from the manufacturing sector, but rather from either agriculture or

---

[5] Emalangeni refers to the local currency unit.

services. The latter appears at face value to ring true as a potential alternative for Swaziland. However, given Swaziland's level of development, its economic development driven by the services sector would constitute premature de-industrialization that might render the country's economic growth trajectory unsustainable at best or divergent at worst; see Rodrik (2015). A recent empirical finding by Rodrik (2015) reaffirms that countries have generally developed a hump-shaped relationship between incomes and industrialization that has moved closer to the origin. This is interpreted to mean that countries are running out of opportunities for entrepreneurial transformation in manufacturing as argued in Schoar (2010).

In Figure 1.2, an evolution of employment by sector is presented for the 10-year period. The number of workers employed in the services sector in Swaziland is on average higher than in the agricultural and manufacturing sectors every year. The agricultural sector started with higher employment relative to manufacturing in the first four years and rose again in the last two years. While the services sector shows an increase in employment in the first four years, it drops in 1998 and starts rising again thereafter.

Figure 1.2: Employment by Sector in the Swazi Economy
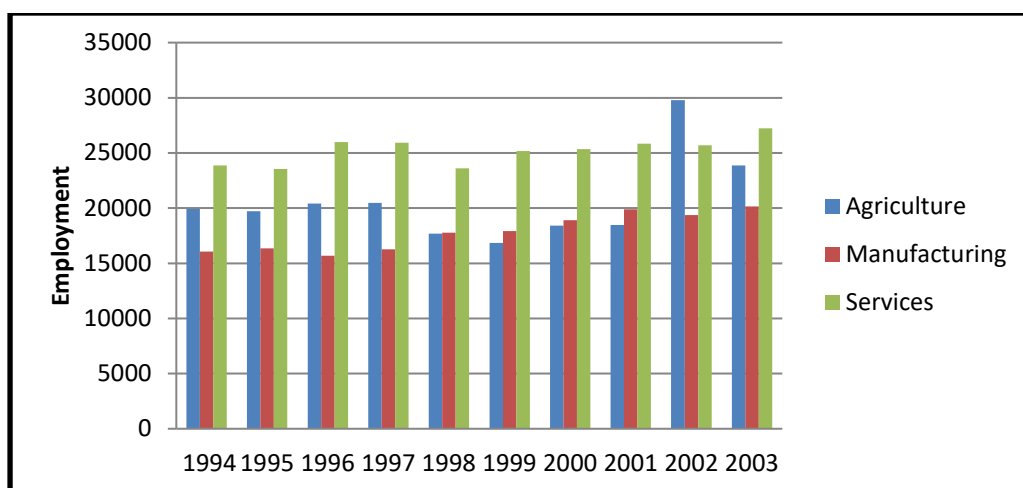
Source: IMF Country Reports 99/13, 00/113 and 06/109.

Figure 1.3 plots annual output/labour ratios to represent sectoral aggregate productivities. The manufacturing sector is on average 3.7 times more productive than the traditional sector, while it is over 1.2 times more productive than the services sector. The services sector is approximately three times more productive every year than the traditional sector. In terms of productivity patterns over time, although performing better than the other sectors; the manufacturing sector experienced persistent deterioration of productivity since 1997 while the services sector shows an improvement since a year earlier.

Figure 1.3: Sectoral Ratio of Real Value-Added to Labour in Swaziland
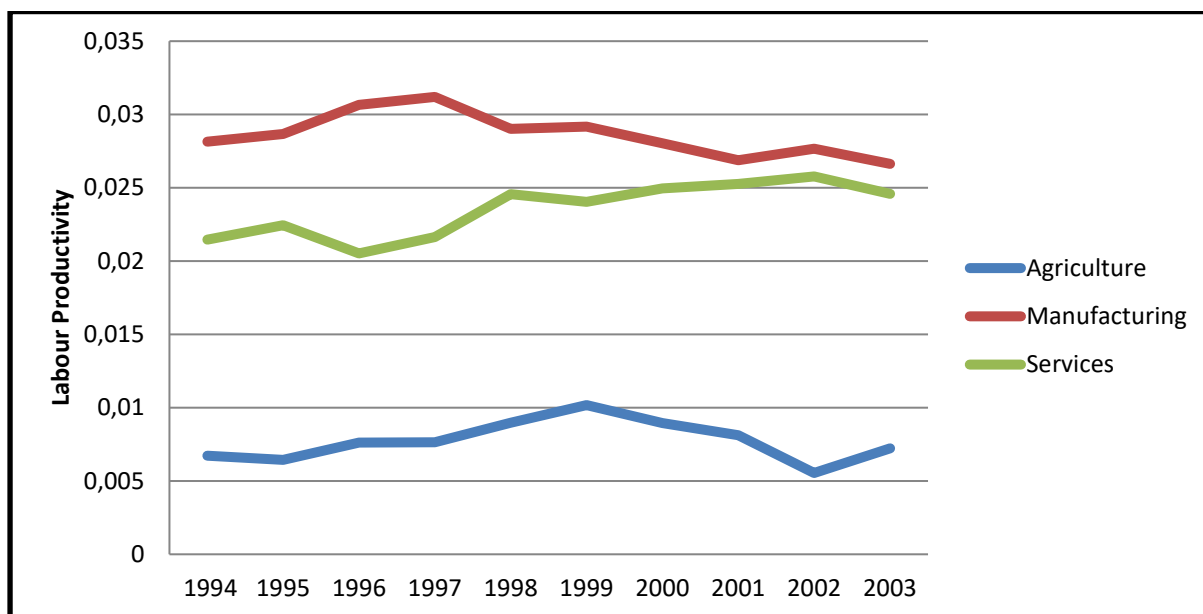
Source: IMF Country Reports 99/13, 00/113 and 06/109.

Taken as a whole, the overall performance of the Swazi economy has been rather mediocre during the decade under investigation. All three sectors basically stagnated, at best, or deteriorated, at worst. The analysis of economic structural change using changes in the country's GDP, however, conceals the underlying microeconomic dynamics that ultimately translate into these macroeconomic outcomes. In the next sections, particular attention is paid to the nature of firm-level data that are used to analyse the behavioural patterns of primary inputs, firm entry-exit dynamics and their individual impact on aggregate productivity growth in the manufacturing sector.

1.3   Panel Data Source and Preparation

The data used in this thesis are constructed from the firm-level "census" data collected by the CSO, which is sanctioned by legislation. Although the survey is supposed to be a census, firms are not specifically targeted unless they contribute a significant amount of output to the sector. A firm may remain off the radar of the CSO until this requirement is met. Furthermore, the response rate also falls short of 100 percent, but firms known to contribute substantially to GDP are followed up until returns are made. This approach has the effect of leaving out of the census a large number of informal and formal microenterprises as well as small manufacturers. Therefore, the probability of an establishment responding to the survey instrument increases with establishment size. Given this data characteristic, the establishment size distribution is likely to be similar to other datasets which have been used for this type of analysis, such as the COMPUSTAT and Ghana, see Axtell (2001) and Sandefur (2010), respectively.

Our dataset contains information on all surveyed manufacturing producers that responded to the questionnaire. This consists of two data-files: the first one includes the component of plant-level output consumed domestically and abroad, employment, wages and benefits, material inputs, energy (combining electricity, water and fuel), and other balance sheet information. The second data-file contains detailed records on each individual establishment's investment and gross capital formation. These include the procurement of and expenditure on land, plant, machinery and equipment, vehicles and other transport equipment, office furniture and equipment.

In order to merge the two data-files, the dataset was cleaned up first.[6] The process of data linkage relied heavily on Christen (2012) to ensure good data quality. If any of the three fields; namely, establishment ID, year or the four-digit ISIC code was empty, the whole record was excluded. The merging of the two files produced a single file with a total of 2 179 records and 335 establishments that ever operated between 1994 and 2003 with identifiable patterns of industrial and export market entry and exit. The resulting dataset was compared with the dataset that was manually compiled by the CSO and also the data published in the World Bank Indicators to establish representativeness. Any differences were accounted for by the inclusion of mining and quarrying establishments and establishments that had either zero or missing values for output, material and/or employees. During the post-2003 period, the CSO experienced technical difficulties with capturing some of its returns, and our data set shows this by the acute decline of establishment count from a total of 171 in 2003 to 128 in 2004 and only 89 in 2010.

To characterize firm entry-exit dynamics in manufacturing, an entrant firm is one that sufficiently expands output and contributes to the top 90 percent of GDP in its industry and enters the database at time t while its ID code is missing at $t - 1$. Firm exit is distinguished by the presence of its ID code at t and a missing ID code at $t + 1$. A continuing firm has its ID code in the database at time $t - 1$ and t for backward calculations or time t and $t + 1$ for forward calculations. The literature favours the first definition of an entrant firm; see Dunne *et al*. (1988). Table 1.1 presents a full schema defining firm entry/exit dynamics adopted in this work.[7]

---

[6] Data quality issues in quantitative research are crucial for the validity of subsequent conclusions drawn. In our case where separate databases are located in different electronic platforms, they need to be combined for ease of analysis. One alternative involves a process of duplicate deletion to ensure a correct history of the firm's performance in the panel is pursued. Achievement of this mission leads to a comparison analysis of variables in the unified database with similar variables in official data to establish representativeness of census data. It is only after these activities and variable definitions that a systematic data analysis is carried out to answer predetermined questions, see Christen (2012).

[7] This definition is similar to Jarmin and Miranda (2002) and Roberts and Tybout (1996) for the Longitudinal Research Database in the U.S.

Table 1.1: The Schema of Firm Entry and Exit Dynamics

Firm Type	$(t-1)$	(t)	$(t+1)$
Entry	Missing	Active	–
Exit	–	Active	Missing
Incumbent-Lag	Active	Active	–
Incumbent-Forward	–	Active	Active

Note: Active means the presence of establishment identity code

The nomenclature reported in Appendix A1.2 is a standardized industrial set of definitions and conventions used in Roberts and Tybout (1996) for developing countries and Haltiwanger *et al.* (2013) for the USA. A detailed specification and robustness checks for capital adjustment based on the perpetual inventory method (PIM) are presented as Appendix A1.3. The capital stock series is robust to small variation in definition and the actual panel data analysis is carried out in the next chapters of this thesis.

There is however a caveat with the panel dataset. We use the U.S. Census Bureau's Longitudinal Business Database (LBD) as a quality benchmark for the micro-level manufacturing data to identify any possible caveat in the Swazi data. Firstly, the data collection instrument makes no provision for distinguishing between a firm, a plant or an establishment. Normally, a firm with multiple establishments receives a unique identity code and its individual establishments are allocated unique identity codes that is not linked to the parent company. As a result, longitudinal linkages that provide for accurate measurement of establishment and firm deaths and births are not available in the dataset. This unavoidably leads to spurious entry and exit dynamics. Furthermore, when a firm or establishment exits the market, it does not retain its original unique identity when it re-enters the industry. Instead, it is issued with a new unique identity code. Again, a change in firm ownership due to either business acquisition or merger does not lead to a change in firm identity code. The purchased firm or establishment simply disappears from the radar of the CSO. This lack of distinction between the firm and its constituent establishments hinders tracking of the dynamics of *both* entities to understand firm growth and entry-exit dynamics. This implies that we can neither calculate between-firm nor between-establishment rates of job flows.[8]

Secondly, the unavailability of information on firm age also prohibits any analysis of the relationship between firm size and net job creation, conditional on firm age. The standard *ad hoc* definition of firm entry based on the first appearance of its unique identity code and using that as a basis for calculating firm age is deficient. If a firm's probability of making it to the radar of the census instrument of data collection is conditional on some administrative criteria, it is probable that the firm may be surveyed

---

[8] If between-firm reallocation rates dominate between-establishment reallocation rates, then such patterns are a reflection of employment shifts between establishments of the same firm, see Davis and Haltiwanger (1999).

years after its actual birth and still be classified as an entrant. As a result, business start-ups and young firms cannot be consistently identified. It is a crucial feature of the data that the majority of small firms, most of which are likely to be young, are excluded from the annual surveys by design.[9] The under-representation and lack of age variable for this group of firms means we cannot reliably assess the relationship between net employment growth and firm size.

1.4 Panel Data Representativeness

The previous section discusses the data sources and implicit quality features in relation to the US benchmark. In this section, the data set is assessed to establish its representativeness of domestic aggregate outcomes. Thus, a high degree of similarity between aggregated firm-level series data and the official or published macroeconomic indicators is treated as evidence of representativeness of the Annual Manufacturing Census of firms. For conciseness, variables of interest are real value added, real capital stock and employment over time and across industry. The closer in magnitude the CSO aggregates are to the published macroeconomic indicators the better. In Table 1.2, these comparisons are made.[10]

Table 1.2: Representativeness of Firm-Level Annual Census Data for the Manufacturing Sector*

	Data from existing macroeconomic sources			Calculated from CSO firm data			
YEAR	Employment (PE)	Capital Stock (E' Million)	Value Added (E' Million)	Employment		Capital Stock (E' Million)	Value Added (E' Million)
	Column 1	Column 2	Column 3	PE	WP	Column 4	Column 5
1994	16 055	828	2 865	16 176	132	903.1	2 414
1995	16 358	1 225	2 979	17 086	150	962.2	2 938
1996	15 969	1 241	3 052	16 396	155	1 081.1	3 122
1997	16 277	1 707	3 219	16 917	95	1 137.7	3 003
1998	17 773	1 978	3 270	18 488	152	1 377.1	3 273
1999	17 905	2 099	3 311	17 907	275	1 666.1	3 421
2000	18 897	1 189	3 360	16 844	307	1 826.7	3 427
2001	19 898	1 129	3 392	26 639	355	1 998.2	4 009
2002	19 370	1 551	3 465	29 879	384	2 157.0	3 420
2003	20 165	1 773	3 527	21 683	307	1 810.9	2 345

Source: Official Macroeconomic Indicators in Columns 1-2 come from the IMF Annual Country Reports (1999, 2000, 2003 and 2006). The real value-added series in Column 3 comes from the World Bank Indicators. Columns for PE and WP as well as Columns 4-5 come from the Annual Census collected by the CSO. PE and WP denote Paid Employees and Working Proprietors, respectively. *Note: *Value added and capital stocks are expressed in constant year-2000 prices.

---

[9] This data weakness prohibits consistent investigations of the role of entrepreneurship on job creation and economic dynamism and prevents determining the relative dominance of subsistence versus transformational entrepreneurship in the manufacturing sector in Swaziland, cf. Decker *et al*. (2014) and Schoar (2010).

[10] An enquiry with the CSO Authorities in July 2016 revealed that published aggregate time series reflect the output of those firms that contribute to the top 90 percent of GDP in the relevant industry. In order to maintain a clear trend, trend smoothing techniques to remove any discernible volatility are then applied. Aggregate datasets submitted to multilateral agencies are subjected to another set of standardization rules for ease of international comparisons.

It is instructive at this point to explain how the two sources of data are compiled. First, published information by the CSO considers large firms that consistently contribute significant output over time. These macro-indicators are also initially standardized by subjecting the time series to trend smoothing techniques to remove any cyclicality, seasonality and any irregularity that may characterize the data. Therefore, any exogenous shocks to the macro system might be muted in the published aggregates. Second, the firm-level panel dataset used has been cleaned up to remove only observations with missing sales revenue and/or employment. In order to replicate the orders of magnitude of the macro-indicators, only paid employment for firms with more than 50 workers is reported, although in subsequent chapters firms employing fewer than 50 workers are used in the analysis. The labour series closely mimic the published aggregates, except for the labour sizes in 2001-2002. In the public data, these two years are likely to have been smoothed out by the Authorities. In the case of real capital stock, while using the perpetual inventory method (PIM) explained in the appendix, the rentals from buildings were not capitalized and a backward calculation was then performed.

For ease of comparison, the employment column in the CSO data separates Working Proprietors (WP) from Paid Employees (PE). The PE column representing employment numbers collected from official macroeconomic sources is compared with employment numbers in the PE column calculated from the CSO panel data set. Although the two series commove in synchrony, the CSO aggregate employment numbers overshoot the official employment numbers in 2001 and 2002. This could potentially be explained as an outcome of strict smoothing procedures implemented by the authorities on official statistics to avoid extreme values of aggregate employment. The columns for the official and panel data real capital stock match as reasonably as could be expected, except for intermittent deviations from one another. Similarly, the real value-added series are well-matched. That is, the real value added series also mimic the macroeconomic indicators when only data from larger firms are considered.

Looking at the firm-level cross-sectional panel data, the manufacturing sector is driven by the performance of only a few tradable commodities in the food, textile, wearing apparel, wood, and pulp and paper industries, see Edwards *et al*. (2013). Table 1.4 presents aggregate statistics compiled from the panel dataset by industry. The columns report proportions of real values of production and capital stock as well as employment for each industry during the 10-year period. The food industry accounts for 19.47 percent of real value added whereas the pulp and paper industry accounts for 47.78 percent. The former is the most labour intensive industry hiring 47.66 percent of manufacturing workers and the latter is responsible for 11.55 percent. However, the pulp and paper industry contains 43.38 percent of the total capital stock in the manufacturing sector.

Table 1.3: Real Value Added and Primary Inputs by Two-Digit ISIC Industry (1994-2003)

INDUSTRY	VALUE ADDED (Percentage)	CAPITAL (Percentage)	EMPLOYMENT (Percentage)
Food (15)	19.47	9.50	47.66
Textile (17)	5.90	2.08	15.83
Apparel (18)	0.74	2.28	2.70
Wood (20)	2.96	2.92	4.16
Pulp & Paper (21)	47.78	43.38	11.55
Printing & Publishing (22)	1.09	2.69	2.29
Chemicals (24)	6.63	5.62	3.28
Rubber (25)	1.35	12.33	0.79
Non-Metallic Mineral (26)	1.41	3.68	2.03
Basic Metals (27)	0.25	2.61	0.23
Fabricated Metal (28)	0.82	3.86	2.21
Furniture (29)	5.38	3.38	2.76
Other Manufacturing (36)	6.21	5.67	4.50

Source: Own calculations from SCO data.

Given the limited domestic market size, producers in each industry tend to focus only on a limited product mix. Local exporters are themselves highly concentrated, such that either a strategic action of one large producer or its vulnerability to a significant external shock can shake up the performance of the entire industry. This was observed in the merger and acquisition involving two firms in the pulp and wood industries in 1998 and can be seen in the underlying data. Edwards *et al*. (2013) note that most manufactures are exported to protected markets in the Customs Union, the European Union for sugar, the U.S. for textile and apparel through AGOA, and Norway for beef through the SACU-EFTA (European Free Trade Area). This fact alone exposes the sector to the risk of preferential treatment erosion with potentially adverse effects on the individual producers, the industry, upstream customers and suppliers of intermediate inputs.[11] Furthermore, commodities traded in the free world markets are subject to price volatility as well as to the Prebisch-Singer thesis, which suggests primary products are likely to experience long-term price deterioration relative to manufactures.

1.5   Conclusion

The analysis of structural change in Swaziland shows a persistent weakening of the manufacturing sector in terms of its share of economic activity and employment relative to the services sector. The manufacturing sector's share of GDP is trending downwards while the agricultural share of output is largely fixed at the same level throughout the period of analysis. During the same period, the services sector's share is trending upwards. The size of the manufacturing sector in Swaziland, the lack of robust industrialization and the limited diversification into globally competitive industrial investments are potential constraints to structural change. In the large and growing literature, the observation is that economic development in Africa is not likely to come from the manufacturing sector, but rather

---

[11] Cf. the deposition by Mulally (2008, pp. 31-32) to the US Committee on the Automobile Industry in Detroit. Also see Gabaix (2011) and Acemoglu *et al*. (2012) for a theory of propagation of idiosyncratic micro-shocks that produce aggregate outcomes.

from either agriculture or services. However, given Swaziland's level of development, economic development driven by the services sector would constitute premature de-industrialization that might render the country's economic growth trajectory unsustainable at best or divergent at worst.

At a micro level, the character of the firm-level data has been evaluated to establish its quality in relation to published macro data. The annual census data for the manufacturing sector in Swaziland closely resembles similar datasets collected by other statistical agencies. In order to analyse the panel dataset directly, the entry-exit dynamics are measured on the basis of a plant's identity code appearing for the first time rather than on firm registration or the last date of existence in the database, respectively. On the whole, we can make the claim that the dataset is of a quality at least as good as any other compiled by a government statistical agency.

The preliminary analysis of the data by two-digit ISIC industry shows the sector's overdependence on a few primary commodities for export to preferential markets. This exposes producers, upstream suppliers of inputs, and downstream customers to potential risk of preferential treatment erosion. For example, a loss of market access for sugar in the EU and U.S. would cause the sugar industry to trade in the volatile world market where sugar prices are generally depressed. Export revenue would decline significantly forcing sugar producers to scale down operations. Likewise, upstream sugarcane farmers would receive reduced revenue, such that the scale of production would also need to be diminished. Again, downstream manufacturers of soft drink concentrates and other users of sugar would have inadequate supplies of this critical input, and may have to import it and incur transport costs. The effect on the whole value chain would be a loss in revenues and employment.

Our future enquiry will take advantage of the newly available and rebased time series on sectoral outputs and inputs to investigate the extent of structural change and the impact of innovation and transformational entrepreneurship 'within' sectors in Swaziland.

APPENDIX

Appendix A1.1: Sectoral Shares of Swazi GDP at 1985 Factor Cost

YEAR	Agriculture	Manufacturing	Services	TOTAL
1994	0.12	0.41	0.47	1.00
1995	0.11	0.42	0.47	1.00
1996	0.13	0.41	0.46	1.00
1997	0.13	0.41	0.46	1.00
1998	0.13	0.41	0.46	1.00
1999	0.13	0.40	0.47	1.00
2000	0.12	0.40	0.48	1.00
2001	0.11	0.40	0.49	1.00
2002	0.12	0.39	0.49	1.00
2003	0.13	0.39	0.49	1.00

Appendix A1.2: Definitions of Variables and Section D Sectors

Variable	Definition
Deflators	Annual deflators sourced from the WBI are weighted by $\vartheta_{it} = \frac{(R_{it}+R_{it-1})}{\Sigma_{i\in\mathcal{E}_{jt}}(R_{it}+R_{it-1})}$, where R_{it} is the establishment i's quantity that is subject to deflation in year t and $\mathcal{E}_{jt}$ is the set of enterprises in the jth four-digit ISIC industry at time t. These deflators include the fixed asset deflator for capital assets, new capital investments, and the manufacturing value added (MVA) for output and material inputs. Wages are deflated with annual inflation sourced from Central Bank Annual Reports.
Output (Q_{it})	The total value of outputs for establishment i in year t to local and foreign markets is $Q_{it} = \frac{DS_{it}}{\vartheta_{it}*MVADEF_t} + \frac{FS_{it}}{\vartheta_{it}*EXPDEF_t} + \frac{NFI_{it}+NWP_{it}}{\vartheta_{it}*MVADEF_t}$, else $Q_{it} = TVS_{it}$ if $NFI_{it}+NWP_{it} = 0$, where $TVS_{it} = \frac{DS_{it}}{\vartheta_{it}*MVADEF_t} + \frac{FS_{it}}{\vartheta_{it}*EXPDEF_t}$ is the total value of outputs for establishment i in year t to local and foreign markets, NFI_{it} is the difference between the values of end-of-year and beginning-of-year finished goods inventories for establishment i in year t, NWP_{it} is the difference between the values of end-of-year and beginning of-year work-in-progress inventories for establishment i in year t, and $MVADEF_t$ is the annual manufacturing value added deflator. When components of NFI or NWP are missing, they are set to zero.
Capital Stock (K_{ijt}^k)	$K_{ijt}^k = I_{ijt}^k + (1-\delta^k)K_{ijt-1}^k$, where I_{ijt}^k denoted the real flow of new investment of asset type k for establishment i in industry j and year t, δ^k represented the rate of depreciation of asset class k, and K_{ijt-1}^k was the previous year's capital stock in industry j. The j subscript is used to refer to a wave in Chapter 4. The capital series is constructed using the perpetual inventory method (PIM).
Labour (Lab_{it})	Use paid employees and working proprietors.
Material (M_{it})	The cost of real establishment-level non-energy material, M_{it}, is calculated as $$M_{it} = \frac{CM_{it}+CS_{it}+CW_{it}}{\vartheta_{it}*MVADEF_t},$$ where CM_{it} is the cost of material for establishment i in year t, CS_{it} is the cost of re-sale products for establishment i in year t, and CW_{it} is the cost of work done for the establishment by others on the establishment's materials in year t. We calculate real energy-water costs as shown in $$E_{it} = \frac{EFW_{it}}{\vartheta_{it}*EPI_{lt}}$$ where EFW_{it} is the cost of purchased electricity-fuel-water for establishment i at year t and EPI_{lt} is the industry-level energy deflator.
Wage (W_{it})	The real wage, $W_{it} = \frac{Wages_{it}}{\vartheta_{it}*CPI_t}$, is the total annual payroll for establishment i at year t, where CPI_t is the Consumer Price Index. This excludes all personnel overheads like housing, transport, pension and others.
Live Sectors	Food and Food Products (15), Textiles (17), Wearing Apparel (18), Wood and Wood Products (20), Pulp, Paper and Paper Products (21), Publishing and Printing (22), Chemicals (24), Rubber and Plastic Products (25), Other non-Metallic Mineral Products (26), Basic Metals (27), Fabricated Metal Products (28), Machinery and Equipment (not elsewhere classified; i.e., n.e.c.) (29), Furniture and Other Manufacturing (n.e.c.) (36)

Appendix A1.3:   **Construction of a Capital Series**

The definition of establishment or firm-level capital in the literature is diverse and ranges from a combination of equipment, machinery, plant, and transport equipment, depending on data availability. In what follows, we use the perpetual inventory method (PIM) to construct capital stocks based on net fixed asset expenditure on plant, machinery and equipment. Our measure of capital stock allows us to initiate the process with the cost of fixed assets per capita and also implement a backward calculation of the series. The net investment, $I_{i,t} = E_{i,t} - Ret_{i,t}$, where $E_{i,t}$ is investment expenditure and $Ret_{i,t}$ is fixed capital retirement, enables us to study the fixed capital asset behaviour. Therefore, we proceed as follows:

Case I: $K_{it} = \delta K_{it-1} + I_{it}$, assuming $I_{it} \neq 0$ for some $i \in n$ and $t \in T$. When this accounting identity produces $K_{it} = 0$ for some $i \in n$ and $t \in T$, then the nominal fixed-capital-assets (z_{it}^{nom}) *plus* capitalized rentals for buildings (z_{it}^{rent}) is used as capital stock. Rent capitalization is partly an outcome of the Government's factory shell programme since the early 1990s.

Sensitivity Analysis

The prevalence of zeros in the establishment-level net investment data set requires careful analysis of the capital stock series constructed. Thus, in order to perform a sensitivity analysis, several assumptions are made and the capital stock series is recalculated to check its robustness. The forward calculation is performed by initializing it with the expenditure on fixed assets and results stored. We then reverse the capital stock computation by starting from the last year and these results are also stored. Both sets of results are then compared with the results in Case I to see if the series is replicated.

Case II: $K_{it} = \delta K_{it-1}$, assuming $I_{it} = 0$ $\forall$ $i \in n$ and $t \in T$. When $K_{it} = 0$, then the real fixed-capital-assets (z_{it}^{real}) *plus* z_{it}^{rent} is used as capital stock.

Case III: $K_{it} = \delta K_{it-1} + I_{it}$, assuming $I_{it} \neq 0$ for some $i \in n$ and $t \in T$. When $K_{it} = 0$ for some $i \in n$ and $t \in T$, then the z_{it}^{real} *plus* $z_{it}^{air} = \frac{1}{3}\sum_{t=-3}^{t}\frac{I_{it}}{K_{it-1}}$, the average investment rate calculated over the three year period prior to $K_{it} = 0$ for some $i \in n$ and $t \in T$, is used as capital stock.

Case IV: $K_{it} = \delta K_{it-1} + I_{it}$, assuming $I_{it} \neq 0$ for some $i \in n$ and $t \in T$. When $K_{it} = 0$ for some $i \in n$ and $t \in T$, then the capitalized depreciation charge $(\frac{z_{it}^{depr}}{0.9})$ *plus* z_{it}^{rent} is used as capital stock.

Case V: $K_{it} = \delta K_{it-1} + I_{it}$, assuming $I_{it} \neq 0$ for some $i \in n$ and $t \in T$. When $K_{it} = 0$ for some $i \in n$ and $t \in T$, then the $(\frac{z_{it}^{depr}}{0.9} + z_{it}^{rent})$ is used. If $\frac{z_{it}^{depr}}{0.9} = 0$ is for some $i \in n$ and $t \in T$, then $(z_{it}^{nom} + z_{it}^{rent})$ is used as capital stock.

Case VI: $K_{it} = \delta K_{it-1}$, assuming $I_{it} = 0$ $\forall i \in n$ and $t \in T$. When $K_{it} = 0$ for some $i \in n$ and $t \in T$, then the $\frac{z_{it}^{depr}}{0.9} + z_{it}^{rent}$ is used. If $\frac{z_{it}^{depr}}{0.9} = 0$ for some $i \in n$ and $t \in T$, then $(z_{it}^{nom} + z_{it}^{rent})$ is used as capital stock.

CHAPTER 2: Job Creation and Destruction in Swazi Manufacturing

2.1   Introduction

Economic development and growth in Swaziland evolved from a period of impressive performance in the 1980s due to incoming resources from the larger and inward-oriented trading partner, South Africa. High-productivity firms and subsidiaries relocated to Swaziland to extract rents from cheaper intermediate inputs and also gain access to foreign markets available to Swazi exporters. As a result, the national gross domestic product (GDP) grew by 11 percent in the manufacturing sector. However, the 1994 political dispensation in South Africa altered the business terrain in the Customs Union, with South African trade liberalization becoming a *de facto* economic policy reform for all member States.

The 1990s therefore marked the beginning of a major turning point in the history of economic development and growth in the Swazi economy. In response to the new competitive pressures, Swazi firms either engaged in structural adjustments to operate lean businesses, or relocated to larger markets including South Africa, or even liquidated their assets and closed down. These firm-level outcomes led to the shedding of labour and lost the ability to employ new workers. Such activities raised the economy-wide level of unemployment to 22 percent in 1995, and by 2007 the level of unemployment had increased to 26.3 percent, with the population aged between 15-24 years hardest hit at 53 percent, see Edwards *et al*. (2013) and Brixiová and Kangoye (2013). Are the observed developments explained by tariff reduction or by South African multinationals with subsidiaries in Swaziland consolidating resources to enhance their own export market shares?

The purpose of this chapter is to investigate the patterns of job flows and to link these to cross-sectional aggregate productivity growth differences during the economic reform period. This is the first systematic evidence of job creation and destruction in the Swazi manufacturing sector. It assesses the evolution of firm-size distribution to confirm or refute the apparent lognormal distribution suggested by the national survey of small and medium enterprises conducted in 2003. The estimation of job creation and destruction makes a distinction between the contributions of small and large firms. There is already a popular belief across economies of the world that small firms are job creators, and this perception is supported by many case studies including Davis *et al*. (1996) for the U.S., provided firm age is not controlled for.

In response to this increasingly robust perception across economies, policymakers in Swaziland formulated a four-fold SME policy that included, *inter alia*, fostering economic growth and development as well as increasing employment opportunities. A set of policy interventions such as access to business finance and training services to support SMEs was adopted, see SME Report (2003). The State also launched a programme designed to localise the new world economic order on

the job creation prowess of small businesses in partnership with Malaysian private sector leading enterprises, as detailed in various reports of the country's Smart Partnership Secretariat.

An important consideration concerns the distinction between subsistence and transformational entrepreneurship that identifies entrepreneurs with the potential for business growth. Schoar (2010) argues that the two types of business owners are not only dissimilar in nature but that an insignificant percentage of them transition from subsistence to transformational entrepreneurship. The majority of entrepreneurs start businesses without the express purpose to innovate and grow in any observable dimension, see Hurst and Pugsley (2011). In the national survey of SMEs in 2003 in Swaziland, only 6.5 percent of the 70 000 small business owners surveyed had an intention to expand their businesses while 28 percent desired only to supplement their income. This suggests a potentially high level of marginal utility of consumption for a large number of small business owners. In respect of creditworthiness, 87 percent applicants were declined credit and 78.1 percent started businesses with their personal savings. Although the educational heterogeneity of owners was pronounced, as much as 77 percent of them had no more than secondary education. In this scenario, an entrepreneur with a suitable mix of business skills, drive for innovation, access to capital and a vision for business expansion is unlikely to emerge. This chapter investigates the creation and destruction of jobs in the Swazi manufacturing sector in order to interrogate these views in more detail.

The organization of the chapter is as follows: the next section presents a review of the literature while section 2.3 reviews the data, identifies any potential caveats in the data and conducts preliminary descriptive analyses of the dataset to extract stylized facts. Special attention is given to the assessment of the evolution of firm-size distribution to establish if the "missing middle" phenomenon associated with developing countries does exist. Section 2.4 addresses theoretical and measurement issues for job creation and destruction. Section 2.5 decomposes the aggregate labour productivity growth into several components and estimates it using firm-level manufacturing data to identify specific drivers of growth in the sector. The last section concludes the chapter.

2.2   Literature Review: Job Creation and Destruction

The general acceptance of small firms' ability to create relatively more jobs than their larger counterparts is reviewed in Davis and Haltiwanger (1996a) and Haltiwanger *et al.* (2013). This hypothesis has become a remarkable regularity in public discourse to justify specific policies in economies such as the U.S., see Decker *et al.* (2016) and Haltiwanger *et al*. (2016). This perception has its origins in the empirical contributions of Birch (1987) that prompted researchers to examine his techniques and to test his assertions in different countries. The data structures in subsequent case studies that support Birch's results suffered from a host of limitations such as lack of data suitability for the purpose, inappropriate firm/establishment size classification, regression to the mean problems,

and failure to draw a distinction between gross and net job creation. Other crucial data constraints involved lack of sample representativeness, and inability to distinguish ownership transfers from birth to death – see Davis *et al*. (1996). Haltiwanger *et al*. (2013) highlight the statistical and measurement complications associated with methods of firm-size classification and the regression-to-the-mean fallacy as major pitfalls in the evidence.[12]

The point of methodological contention with Birch's work is based on his measure of firm growth using the difference between time t and time $t-1$ employment divided by the base year employment at $t-1$. As argued in Davis *et al*. (1996a), this introduces the regression-to-the-mean problem. Several sources of this problem have been advanced. For example, a business experiencing a negative transitory shock at $t-1$ or subject to a transitory measurement error is likely to grow at t, while a business that experiences a positive temporary shock at $t-1$ is likely to contract at t, see Davis *et al*. (1996a, b). On the same basis, Friedman (1992) concludes that this type of regression fallacy "is the most common fallacy in the statistical analysis of economic data". Davis *et al*. (1996a, b) propose and popularize a firm-size classification methodology that relies on the average of the base year $(t-1)$ and current year (t) to mitigate the effects of the regression fallacy.

However, the Davis *et al*. (1996a, b) approach is also vulnerable to effects of permanent shocks that move the plant across multiple size-class boundaries between $t-1$ and t. The plant then becomes categorized in a size class that is between the starting and ending size classes (Davidsson, Lindmark and Olofsson, 1998). In spite of this limitation, however, its results are robust to the kind of dynamics introduced by Butani, Clayton, Kapani, Spletzer, Talan, and Werking (2006) which attributes job gains or losses to each size class that the firm passes through, see Haltiwanger *et al*. (2013). The technical concerns highlighted by Davis *et al*. (1996a, b) still do not obviate the perceived ability of small firms to create the most jobs in empirical work, including in developing economies.

In the latter literature represented by Haltiwanger *et al*. (2013), the role of firm age is shown to have a real impact on the job creation ability of small firms. It finds new births and young firms that happen to be small, to create the most jobs. Using an enhanced U.S. Longitudinal Business Database, Haltiwanger *et al*. (2016) find that high-growth young firms contribute disproportionately to job creation, and they identify high heterogeneity among young firms in terms of the failure rate in the first few years of existence. The growth patterns among surviving firms are characterized by marked dispersions. Conditional on survival, young firms have higher average net employment growth compared to their more mature counterparts. In the U.S. case, firm-level net employment growth is characterized by robust positive skewness with a small fraction of very fast growing firms driving the

---

[12] However, Davidsson, Lindmark and Olofsson (1998) use Swedish data to estimate the extent to which job creation by small firms is overestimated by using Birch methods and find insignificant regression bias effects.

higher average growth of net employment. This is a typical case of small firms' transition to transformational entrepreneurship.

Using a good quality firm-level manufacturing panel dataset, our investigation performs the first systematic job flow analysis in Swaziland covering the period 1994-2003. Access to such data provides a valuable opportunity to understand the main sources of employment variation. Specifically, we slice firms into small and large class sizes of less and more than 50 workers, respectively. For each firm-size category, different types of job flows are calculated over time and across two-digit ISIC industries. This allows us to answer questions about sources of job flows by firm size over time and across two-digit ISIC industries.

2.3   The Data and Descriptive Analysis

The Central Statistical Office of Swaziland conducts annual census data collection for firms that contribute to the top 90 percent of the industry's gross domestic product (GDP). Although this is treated as a census of firms, there is still a non-negligible level of nonresponse by some producers. The dataset does not distinguish between a firm, a plant or an establishment. These terms are used interchangeably here to refer to the unit of observation. Each firm-year observation is identified by a unique time-invariant identity code. It is therefore possible to track a business unit from "start-up" stage when it appears in the data for the first time to its exit, at least in theory. However, sometimes a firm is observed at entry, then disappears for a year and reappears in the data. When this occurs, it is interpreted as a case of nonresponse or administrative measurement error rather than the firm shutting down for a year and resurfacing to produce the same product.

2.3.1   Descriptive Statistics

This section looks at descriptive patterns of establishment size distribution in the context of popular belief that producer size dynamics are scale dependent; that is, small manufacturers grow faster than larger ones conditional on survival, see Rossi-Hansberg and Wright (2007). Such dynamics should necessarily feature prominently in patterns of plant size transition probabilities where a significant proportion of small firms at time t would cross-over to a larger size category at time $t + 1$. This notion has, however; been challenged by evidence in Sub-Saharan Africa through identifying the presence of a dense mass of smaller firms coexisting with fewer medium sized establishments together with even fewer large firms, see Gebreeyesus (2008). This line of enquiry is complemented with investigations of the entry-exit dynamics to understand the degree of firm churning during periods of trade reforms. The descriptive analysis further focuses on an elaborate analysis of the "Missing Middle" hypothesis to determine its validity or otherwise in the case of the Swazi manufacturing sector. It concludes with studying the distribution of employment by industry and plant size category as well as employment growth along similar dimensions.

Figure 2.1 shows the pattern of establishment distribution over a ten-year period. This is a period of trade liberalization in the customs union that also experienced industrial development of factory shells in 2001. Analysing the population of firms sliced into two size categories: small firms employing at most ($\leq$) 50 workers and large firms employing more than ($>$) 50 workers, allows us to detect key firm size dynamics. As expected, the mass of smaller firms constitutes approximately two-thirds of the population of surveyed manufacturing producers in any one year. In this case, the population growth rate is likely to be firm-size independent. This scenario can arise for at least two reasons. First, it may occur when the entry-exit dynamics are the same across firm sizes. Second, it may also be that the rate of firm size growth is the same across firm size categories; that is, transition probabilities from small to large firms and vice versa are the same.

Figure 2.1: Annual Number of Firms by Firm Size Category
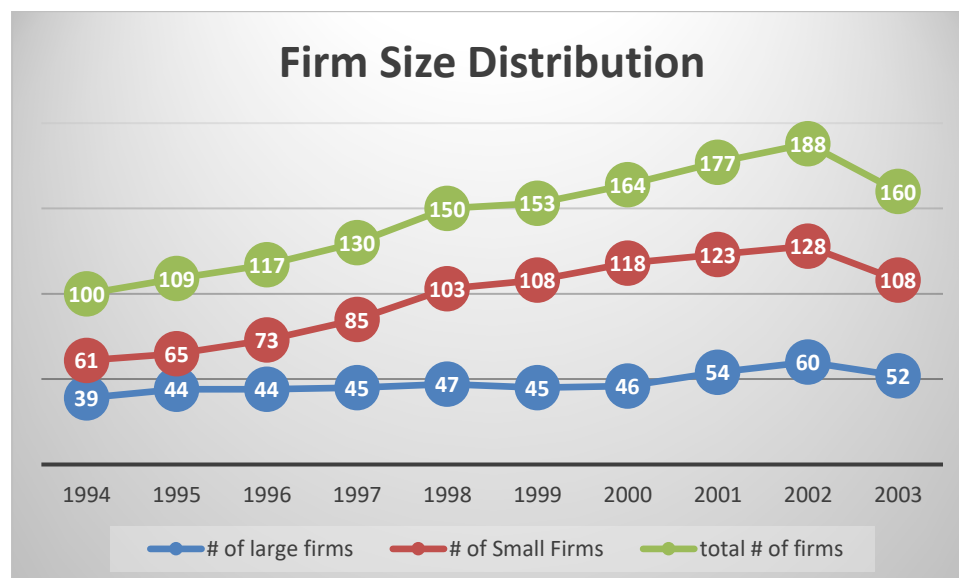
Source: Author's calculations from data compiled by the CSO

Table 2.1 therefore presents a matrix of firm size transition probabilities. Slicing the panel dataset of establishments into different size categories allows us to understand cross-over rates of firms between groups. One way to perform these calculations is to assume that there are missing values of firms so that the data can be rectangularized for normalization in order to compute Markov transitions. Another way is to simply order the data by time *t* and count transitions of observations between states. The latter approach is used here to produce our results.

Table 2.1: Firm-Size Transition Probabilities

	Firm Size		
Firm Size	0	1	Total
0	98.02	1.98	100.00
1	1.72	98.28	100.00
Total	33.66	66.34	100.00

Source: Author's calculations from data compiled by the CSO

Thus, using 0 as a proxy for firms employing ≤50 workers and 1 as a proxy for firms employing >50 workers in manufacturing, the table shows that each year, 98.02 percent of small firms remained small while the 1.98 percent of small firms transitioned across the 50 employee threshold to enter the large firm category. At the same time, large firms had a 1.72 percent probability of crossing over the size category to become smaller firms. These results remain robust to using Markov transition estimation methods as in Sandefur (2010). Therefore, large firms in Swazi manufacturing were born large and did not grow out of the small size category while small plants remained small.

2.3.1.1 Entry and Exit of Firms in Manufacturing (1994-2003)

The schema of firm entry and exit presented in Table 1.1 provides measurement methods for firm churning dynamics. This section begins with the measurement of four-digit ISIC industries' average results of firm entry and exit as well as the churning rate in each year in Table 2.3. As expected from the firm-size distribution results, entry rates exceed exit rates during the 1995-2001 period and exit rates become spiky in 2002 and 2003. These spikes are a reflection of the general disappearance of firms from the sample since 2003 due possibly to non-response to questionnaires or non-digitization of responses from establishments, or a combination of factors. The churning rate (sum of entry and exit rates) shows considerable volatility ranging from 9.26 percent in 1999 to 44.06 percent, in 2003, reflecting the simultaneous entry and exit of firms in manufacturing.

Table 2.3: Establishment Churning and Survival Rates in Manufacturing

	1995	1996	1997	1998	1999	2000	2001	2002	2003
Entry	0.0862	0.1339	0.1014	0.1355	0.0864	0.0747	0.0806	0.0808	0.096
Exit	0.0517	0.0236	0.0217	0.0387	0.0062	0.023	0.0269	0.1869	0.3446
Churning	0.1379	0.1575	0.1231	0.1742	0.0926	0.0977	0.1075	0.2677	0.4406
Survivor (−1)	0.9138	0.8661	0.8986	0.8645	0.9136	0.9253	0.9194	0.9192	0.904
Survivor (f)	0.9483	0.9764	0.9783	0.9613	0.9938	0.977	0.9731	0.8131	0.6554

Source: Author's calculations from data compiled by the CSO
Notes:
- Entry is defined as a firm missing at $t − 1$ and active at time t.
- Exit is defined as an active firm at time t and missing at time $t + 1$.
- Lagged survivor means a firm is active at times $t − 1$ and t.
- Forward survivor means a firm is active at times t and $t + 1$.

The measure of firm survival rate is based on the number of firms present at $t-1$, with t calculated using the time lag method or those present at t, and $t+1$ calculated using the forward lead method. As gleaned from the table, a sudden increase in the exit rate leads to a sudden drop in the forward survivor rate since these are inextricably linked. A similar pattern between the entry rate and lagged survivor rate is observed. A deeper scrutiny of these results shows that Entry Rate + Survivor(-1)=100 percent and Exit Rate + Survivor Rate(f)=100 percent. Thus, both definitions of the survival rate produce useful results.

2.3.1.2 Annual Employment by Industry Category

In this section we look at the behaviour of firm-level employment over the 10-year period by ISIC industry in order to get a sense of the growth of firms in each industry. In table 2.4, different industries responded differently to external shocks, with some displaying spiky growth of employment such as observed in the Food (15), Textile (17), Clothing (18) and Wood (20) industries. The Food industry, the leading employer, includes key national export products like sugar, soft drink concentrates, fruits and nuts as indicated in Edwards *et al*. (2013). Other industries either remained marginally unchanged in employment levels or declined like the Pulp and Paper (21), Basic Metal (27) and Furniture (29) industries. In almost all industries, the number of firms declined in 2003.

Table 2.4: Annual Employment by Two-Digit ISIC Industry

ISIC	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
15	7 891	8 259	8 443	9 092	9 671	11 432	11 932	15 214	15 534	11 241
17	2 088	2 081	1 103	1 508	2 513	2 966	3 132	5 574	8 263	5 281
18	28	30	81	85	86	94	95	1 676	1 866	1 874
20	855	932	827	837	894	874	929	870	1 119	957
21	3 351	3 543	3 463	2 827	2 857	2 829	2 024	1 533	1 663	848
22	396	432	494	517	527	492	506	567	466	755
24	611	598	666	633	696	788	852	873	870	624
25	108	118	126	181	182	193	231	245	252	130
26	405	434	477	507	506	445	491	507	494	309
27	157	150	27	20	19	20	36	34	46	35
28	296	268	343	378	385	452	507	811	807	844
29	700	741	919	923	1 029	1 017	191	174	163	151
36	507	783	1 025	1 100	1 084	1 083	1 088	1 138	1 084	950
FIRM SIZE										
Firms≤50	1084	1130	1441	1596	1808	1853	2192	2222	2340	1816
Firms>50	16176	17086	16396	16917	18488	17907	16844	26639	29879	21683
TOTAL	17260	18216	17837	18513	20296	19760	19036	28861	32219	23499

Author's calculations from data compiled by the CSO

Looking at employment levels for small firms shows a monotonic increase in employment until 2002 and a decline in the last year (2003). In contrast, large plants' employment growth oscillates somewhat in the first eight years, spikes in 2001 and 2002, and declines in 2003. The years from 2000 and 2002 appear to be a period of relatively high activity. This size-dependent heterogeneity suggests that there might be fundamental differences between small and large firms in manufacturing.

2.3.1.3 Employment Growth in Manufacturing in Swaziland

A few broad messages are gleaned from Table 2.5 and Table A.2.1 which present results over time and cross-sectionally. Although the median annual firm growth is positive every year but 2003, the average growth rate over time is negative only in those years when industries experience business closures by large firms. That is, the negative average numbers are explained largely by the fall in employment growth at 139 percent in Basic Metals (27) in 1996 and at 137 percent in Furniture (29) in 2000. The large positive mean employment growth rate of 22 percent in 2001 is driven by a positive shock in the Apparel (18) industry due possibly to the AGOA effects. As expected, the highest volatility in employment growth is observed only in the years 1996, 2000, 2001 and 2003. However, the Food (15) industry shows the most significant and robust resilience in employment growth, except only in 2003.

Most employment growth, however; came from in the Apparel and Clothing (18) and Textile industries with an average of 38.13 percent and 17.13 percent, respectively. This occurred in the backdrop of significant volatility of 64.54 and 37.43, respectively. Since the transition probabilities suggest absence of firm-level longitudinal growth or contraction, the growth patterns observed in these industries resulted from entry of new firms. That is, these patterns of firm growth are largely a measure of the impact of entry-exit dynamics observed in the manufacturing sector. For example, while the exit rate ranged between 0.62 percent and 5.17 percent in 1999 and 1995, respectively; it started developing into a spike growing from 18.69 percent in 2002 to 34.46 percent in 2003. Similarly, the highest rate of firm entry was experienced in 1998 at 13.55 percent as a result of mergers and acquisitions in the Wood and Pulp industry. On the other hand, the reduction of protection induced closure of high cost producers mainly in the Basic Metal and Furniture industries due to loss of competitiveness, see Table A.2.2 in the Appendix.

Table 2.5: Annual Employment Growth Rate, $g_{it} = \frac{(N_{it} - N_{it-1})}{(1/2)(N_{it} + N_{it-1})}$, by Two-Digit ISIC Industry

	YEAR									FIRM SIZE	
ISIC CODE	1995	1996	1997	1998	1999	2000	2001	2002	2003	Firms≤50	Firms>50
15	0.05	0.02	0.07	0.06	0.17	0.04	0.24	0.02	-0.32	2.58	0.42
17	0.00	-0.61	0.31	0.50	0.17	0.05	0.56	0.39	-0.44	-0.45	0.36
18	0.07	0.92	0.05	0.01	0.09	0.01	1.79	0.11	0.00	-0.06	0.05
20	0.09	-0.12	0.01	0.07	-0.02	0.06	-0.07	0.25	-0.16	0.22	0.01
21	0.06	-0.02	-0.20	0.01	-0.01	-0.33	-0.28	0.08	-0.65	-0.07	-0.26
22	0.09	0.13	0.05	0.02	-0.07	0.03	0.11	-0.20	0.47	-0.13	0.16
24	-0.02	0.11	-0.05	0.09	0.12	0.08	0.02	0.00	-0.33	-4.50	0.10
25	0.09	0.07	0.36	0.01	0.06	0.18	0.06	0.03	-0.64	0.40	0.07
26	0.07	0.09	0.06	0.00	-0.13	0.10	0.03	-0.03	-0.46	-0.88	-0.10
27	-0.05	-1.39	-0.30	-0.05	0.05	0.57	-0.06	0.30	-0.27	-0.18	-0.02
28	-0.10	0.25	0.10	0.02	0.16	0.11	0.46	0.00	0.04	1.01	0.26
29	0.06	0.21	0.00	0.11	-0.01	-1.37	-0.09	-0.07	-0.08	0.13	0.03
36	0.43	0.27	0.07	-0.01	0.00	0.00	0.04	-0.05	-0.13	-0.55	0.29
Total	**0.84**	**-0.07**	**0.53**	**0.84**	**0.58**	**-0.47**	**2.81**	**0.83**	**-2.97**	**-0.19**	**0.10**
Average	**0.06**	**-0.01**	**0.04**	**0.06**	**0.04**	**-0.04**	**0.22**	**0.06**	**-0.23**	**-0.19**	**0.11**

24

Median	0.06	0.09	0.05	0.02	0.05	0.05	0.04	0.02	-0.27	-0.07	0.07
Std Dev	0.12	0.53	0.17	0.14	0.10	0.44	0.52	0.16	0.30	1.55	0.19

Looking at the industrial growth performance by firm size provides useful information. A distinction between small and large firms by industry enables us to assess the related decomposition of firm-size growth. Firms with less than 50 workers reduced employment by 19 percent on average while larger ones increased their employment by 11 percent every year on average. Put differently, unlike in the US case which displayed a heightened pace of business dynamism and entrepreneurship in the period preceding 2000 in Decker *et al*. (2016), large firms in the Swazi manufacturing were generally born large. Large firms did *not* emerge from a growing mass of small dynamic and entrepreneurial businesses.

A few exceptions of positive growth were however identified at the two-digit industry level. Small establishments in the Food (15) and Fabricated Metal (28) industries experienced significant growth. The latter industry is an important upstream supplier of key inputs to the Textile (17) and the Wearing Apparel (18) industries. Marked plant-level growth among larger firms was observed in the Food, Textile, Printing and Publishing (22) and Fabricated Metals (28). Thus, large firm growth dominates the growth of firms employing less than 50 workers. However, the robust small firm decline is characterized by heterogeneity that is 8.2 times higher than that of larger firms. This suggests the high sensitivity of small firms to exogenous shocks that influenced the performance of business and industry during the period of trade reforms in the Customs Union. The observed divergence in the growth of small and large firms may sound a warning for the development of a bimodal firm-growth distribution in the sector. This hypothesis is formally studied and tested in section 2.3.2.4.

While useful as far as it goes, Table 2.5 is unable to provide direct information about the impact of the entry-exit dynamic on firm-level growth. It also provides no guidance on the incidence of incumbent plants that neither supply new nor reduce existing jobs. The literature, led by Davis and Haltiwanger (1992), suggests that the growth rate of establishments is symmetric about zero and lies in the closed interval [-2, 2], with shut-downs and start-ups respectively corresponding to the left and right endpoints. This measure facilitates a unified treatment of plant entry, exit and continuation in the analysis of employment dynamics. Using Eq. 1, we calculate growth rates of firms by size category and plot the frequency distributions in Figure 2.2. As in Davis and Haltiwanger (1992) for the U.S. and Shiferaw and Bedi (2009) for Ethiopia, these densities are asymmetric with central peaks in the interval around the zero growth rate and endpoint spikes corresponding to firm deaths and births. The firm-growth distribution reveals that firm entry-exit dynamics were important for job creation and destruction during the economic reform period. Larger firms are characterized by higher entry and exit activity than firms employing at most 50 workers, and a relatively greater mass of firm growth is concentrated around the centre with the distribution decaying along the tails. As seen earlier, plant

turnover and employment volatility is a function of sharply declining firm size in manufacturing during the period of trade liberalization.

Figure 2.2: Frequency Density of Job Growth Rates
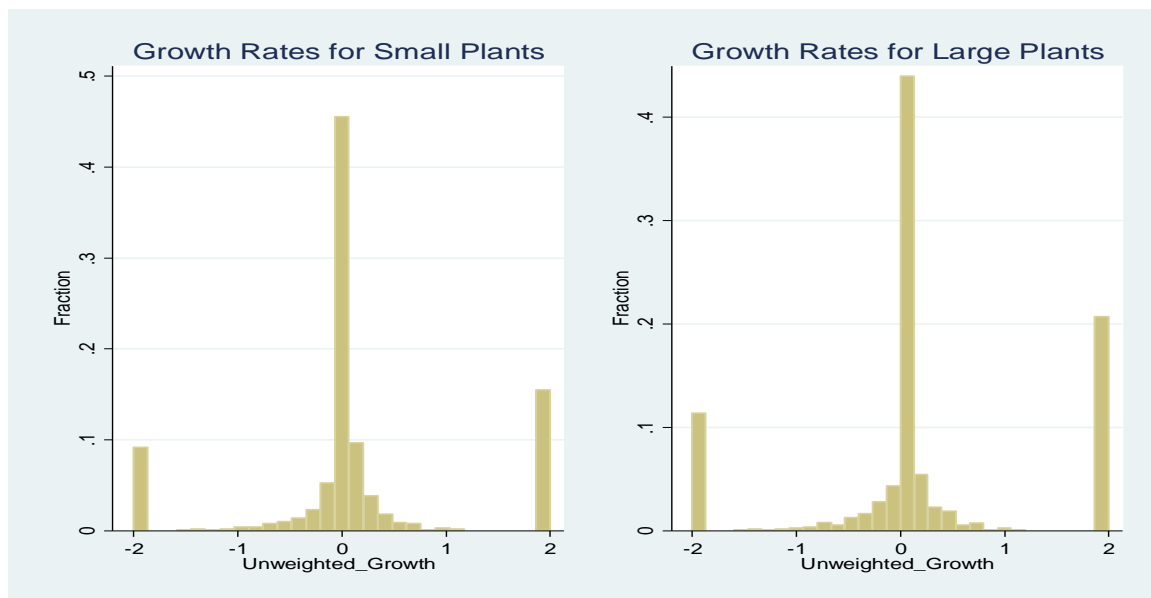
Source: Author's calculations from data compiled by the CSO

2.3.2   Is the "Missing Middle" Hypothesis a Valid Industrial Proposition for Swaziland?

2.3.2.1 Firm Size Distribution

This section considers firm size distribution over the 10-year period to establish any potential presence of the "Missing Middle" phenomenon in manufacturing as claimed to exist in Sub-Saharan African economies. Proponents of its presence in developing countries include Dasgupta (2010) and Mazumbdar and Sarkar (2008). However, this developing conventional wisdom has been challenged by Teal (2016) for Ghana, Hsieh and Olken (2014) for India, Mexico and Indonesia and by Tybout (2014) using normative arguments. In analysing the case of the missing middle in the Swazi data, visual methods for the firm size distribution are initially used and the approach adopted by Hsieh and Olken (2014) is subsequently applied.

Figure 2.2 therefore plots the evolution of the cross-sectional density distribution by firm size to establish if there is more than a single peak in firm-level employment. The firm-size distribution in the first panel is based on selected cross-sections of four years and the second panel is a firm-size distribution of all the years under study. In both panels, there is a clear gradual shift of firm sizes to what visually seems like two clusters or twin peaks. One cluster is around the highest modal value of the natural logarithm (20 employees) and the lower concentration is around the natural logarithm (403 employees). While generally resembling a lognormal distribution, it is interesting that the firm-size

26

distribution *appeared to* evolve gradually from a unimodal structural distribution in 1994 to a bimodal distribution in 2003.[13] It may be that trade liberalization of the 1990s and mid-2000s in the Customs Union resulted in the gradual exit of productive medium-sized foreign subsidiaries to relocate in larger markets. Such self-selection may have created a gap between small unproductive firms and larger productive ones that remained due to, among other reasons, capital irreversibility.[14]

The evolutional structural change in the manufacturing sector in Swaziland resembles that of Ghana's single cross-section in 2003, see Sandefur (2010).[15] It initially exhibits a lognormal distribution that cross-sectionally shifts to the left and gradually develops differential rates of change after the modal point over the 10-year period. More specifically, the cross-sectional distribution of log employment experienced a gradual decline in the rate of change after reaching its modal level; albeit, without a discernible tendency towards a second peak. There are at least three potential explanations for a twin-peak firm size distribution in the development economics literature; see for example Quah (1997), Krueger (2013) and Dasgupta (2010).[16] First, small firms may be financially constrained and unable to grow to become medium sized. The process leads to a "missing middle" situation. Second, a dual economy may develop due to high fixed costs of regulation for large firms, which deters middle-sized firms from growing. Third, the self-selection of agents into either the traditional or modern sector based on the level of individual competitive advantage in knowledge and innovation with a low mean firm size may produce the missing middle feature in manufacturing. Typically, as the mean firm-size increases, the firm-size distribution converges from the bimodal to a unimodal density distribution in the development trajectory of nations. The next sub-sections pursue a rigorous inquiry into whether or not a missing middle exists using returns to primary inputs.

---

[13] Dasgupta's (2015) theory suggests the "missing middle" phenomenon is a stage of development characteristic of developing countries. In the manufacturing data for India, Indonesia and Mexico, Hsieh and Olken (2014) find no evidence of the "missing middle" while Mazumdar and Sarkar (2008) confirm its persistent presence in India and Sandefur (2010) detects its presence for the case of Ghana.

[14] Our definition of the "missing middle" is immune to the Hsieh and Olken (2014) critique based on binning firm size classes. It is also richer than the case of Ghana which focuses on a single cross-section of 2003 firms, see Sandefur (2010).

[15] Hsieh and Olken (2014) present strong arguments against bimodality in developing countries based on manufacturing data from India, Indonesia and Mexico. In their objection, they argue that developing economies have a few main characteristics: missing medium-sized *and* large firms at the same time, higher average product of labour and capital for small firms than for larger ones and absence of material discontinuities in firm-size distribution that may suggest the presence of regulatory obstacles for large firms.

[16] Hsieh and Olken (2014) refute the "missing middle" hypothesis for developing economies arguing that both medium-sized and large firms are missing.
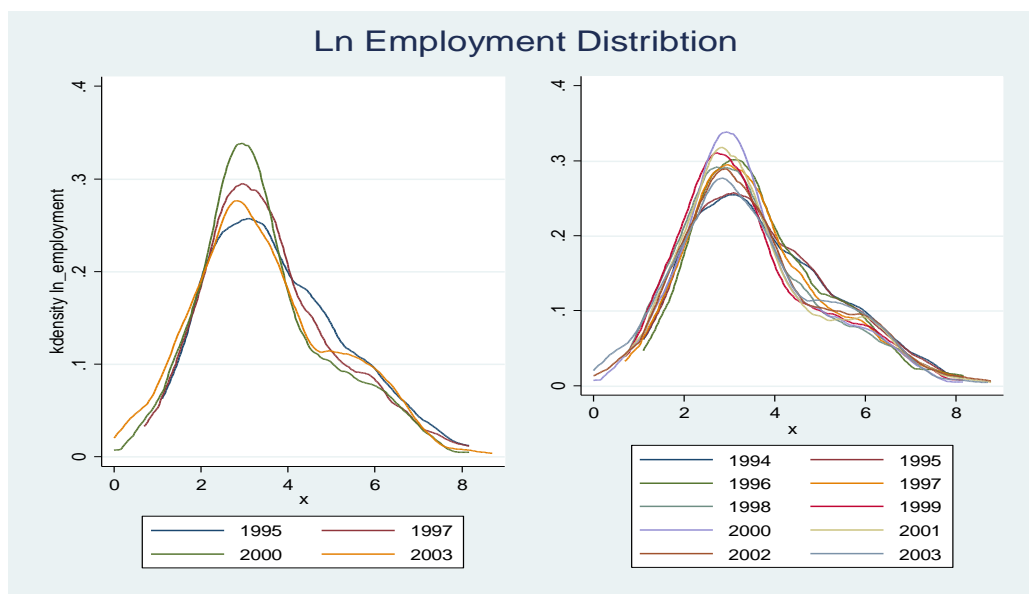
Figure 2.2: Firm-Size Distribution (1994-2003)

Source: Author's calculation based on data from CSO.

2.3.2.1.1   The Distribution of Average Returns to Primary Inputs

As discussed thus far, measurement issues around the definition and the inherent nature of the "missing middle" are a subject of continuing discourse. Key among these is the idea that small enterprises are constrained to grow due to lack of credit capital while medium-sized firms grow to become large, leaving a firm-size gap in the middle of the distribution. On the other hand, there is the 'dual economy' proposition of co-existence of small low-productivity enterprises with large and disfavoured high-productivity firms. This hypothesis argues that the requirement for larger firms to bear large fixed costs of regulation deters the potential growth of medium-sized firms to become established as well. Micro-based country studies such as Hsieh and Olken (2014) on India for 2011, Indonesia for 2006 and Mexico for 2008 find unimodal distributions of average returns to individual and joint primary inputs. This suggests the absence of the missing middle in the referenced economies. In contrast, working on the Vietnamese manufacturing data for the 2000-2008 period, Pham and Takayama (2015) find that the "missing middle" in the distribution of production efficiency is present.

A choice in the definition of what constitutes a missing middle and the related method of analysis appear to drive the results observed in empirical studies. The analysis of the Swazi case relies on the heterogeneity of the marginal product of primary inputs that we proxy with the average product of input due to data availability, see Hsieh and Olken (2014). In the Appendix, Figure A.2.1 graphically looks at correlation patterns between the average product of capital and firm-size. This analysis is

complemented with a more rigorous statistical evaluation of the average product of output relative to real capital stock in Figure A.2.3.

The graph shows a largely flat distribution of the average product of capital in relation to employment. This may be due to the high degree of capital irreversibility and high adjustment cost of the plant and machinery component of capital stock across establishments over time. *Ceteris paribus*, and assuming proportionality of the average to the marginal product of capital, the observed empirical distribution predicts that the marginal cost of capital is largely insensitive to firm size. Such a reality is not consistent with the commonly held notion that the return to capital is higher in small establishments in developing economies. Put differently, if the return to capital in small firms is low, the evidence here suggests that the return to capital in larger firms is not significantly different from that of small firms.

However, the most relevant variables of interest in our context is the study of behavioural patterns of the average product of labour in relation to establishment size. Figure 2.3 plots the nonparametric correlation between the average product of labour and firm size. As can be seen, the relationship is negative. This implies that marginal costs of labour are declining with employment size, in sharp contrast to the Banerjee and Duflo (2005) model of dual technology which posits existence of high capital-intensity for modern technologies. A potential explanation for this pattern is that larger establishments charge lower mark-ups due to economies of scale relative to smaller firms and their exposure to international competition, see De Loecker and Warzynski (2012) and Hsieh and Olken (2014).

Figure 2.3: Average Product of Labour by Employment (1994-2003)

Graphs by Year
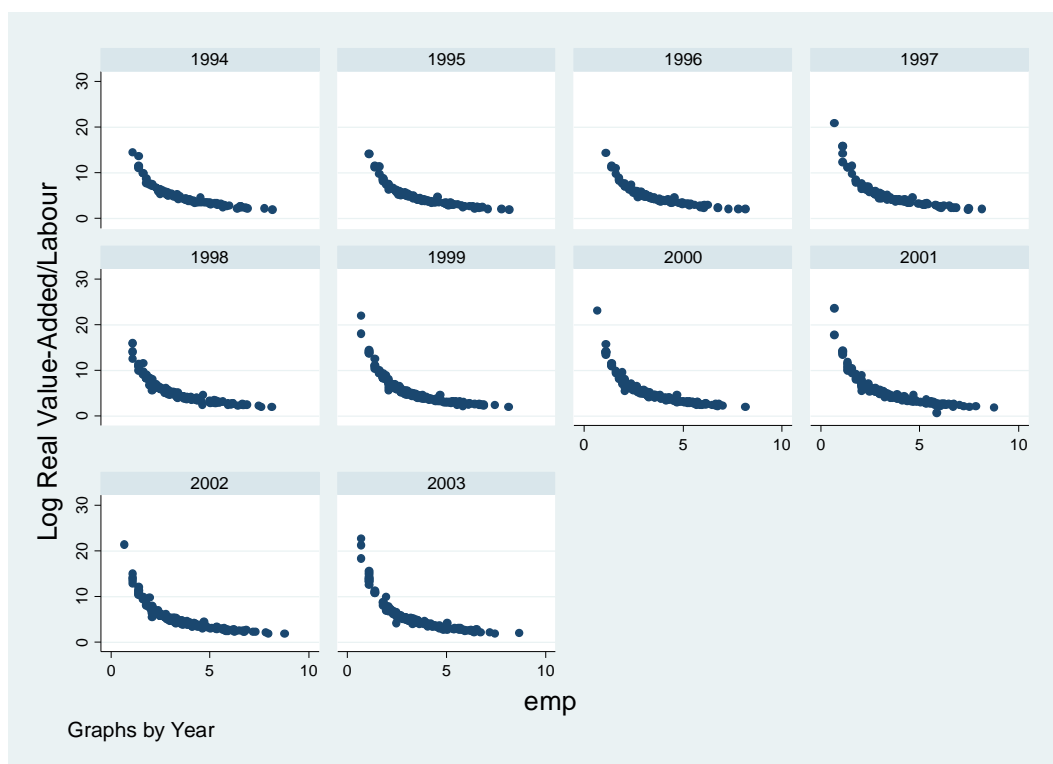
Taken together, Figure 2.3 and Figure A.2.2 produce a stylized fact that there is *no* obvious evidence of bimodality in any of these empirical distributions. However, it remains crucial to produce conclusive results of the unimodality versus multimodality conundrum about the average return to primary inputs using statistical tests.

2.3.2.1.2   *Numerical Test Methods for the Missing "Missing Middle"*

The previous section has visually established the missing bimodality distribution of the average product of capital and labour but has implicitly suggested the need for a rigorous analysis of the bimodality hypothesis in the study of the missing middle. A variety of methods exists in the statistics literature for assessing statistical significance of empirical distributions, but the Hartigan and Hartigan (1985) dip test of unimodality is more appealing and widely used for this purpose. This test is designed to measure the maximum difference between the empirical distribution and the unimodal distribution functions that minimizes this difference. It is calculated in n operations for n observations and its null is that the empirical distribution function is unimodal. In Hartigan and Hartigan (1985), the p-values are computed by matching the dip statistic obtained with those for repeated samples of the same size from a uniform distribution. This procedure also reports extraneous statistics such as the low, high and mean of the modal interval for the best-fitting uniform distribution corresponding to the data, see Hartigan and Hartigan (1985).

In table A.2.1, we report dip test statistics for the average product of capital for all firms in each year and corresponding *p*-values. The dip test shows insignificant statistics every year; hence, failing to reject a unimodal empirical distribution of the average product of capital. Similarly, the average product of labour is tested in Table 2.2 and, again, the dip test results refute any presence of bimodality in the empirical distribution function.

Table 2.2: The Dip Test of the Average Product of Labour

Year	Firms	Dip	*p*-Value	Low	High	Mean
1994	100	0.04	0.26	2.46	4.33	3.41
1995	109	0.04	0.16	3.22	4.09	3.65
1996	117	0.02	0.83	3.69	4.09	3.88
1997	129	0.02	0.80	3.63	4.08	3.84
1998	147	0.02	0.89	4.72	5.20	4.98
1999	152	0.02	0.84	4.50	5.05	4.80
2000	162	0.01	0.99	4.74	5.10	4.93
2001	177	0.01	0.98	4.93	5.09	5.01
2002	185	0.01	0.98	4.07	5.41	4.75
2003	157	0.02	0.49	4.23	5.39	4.82

Clearly, graphical and statistical investigations of bimodal empirical distributions of average products of inputs in manufacturing during the period under study produce distributions characterized by a missing "missing middle".

2.3.3   Stylized Facts

The descriptive analysis produced several stylized facts about firm-size distribution and the missing middle hypothesis in the manufacturing sector. These are:

2.3.3.1 Large Firms are Born Large in Swaziland

The firm size distribution shows that small firms constitutes over 66 percent of the sample population in manufacturing sector. An analysis of the probability of firms transitioning across size boundaries indicates that 98.02 percent of firms born small remain small in the entire period while 98.28 percent of firms born large remain large for the whole period. Thus, the cross-over of incumbent firms from one size category to another through employment adjustments was non-existing.

2.3.3.2 Large Firms in Key Industries Experienced a Jump in 2001 and 2002 Employment.

The net entry of large producers was 8 firms in 2001 and 6 firms in 2002, increasing jobs by 58.15 percent and 12.16 percent, respectively. Significant changes in firm turnover occurred in the Food, Textile and Apparel industries potentially because of their high export propensity.

2.3.3.3 Firm Size Dynamics Exhibit a Lognormal Density Distribution

The cross-sectional firm size lognormal distributions covering the 10-year period shifted towards the left. This distribution evolves overtime to a state where it initially declines rapidly after its modal level, then slows down as if to form a 'dip' before accelerating again. The pattern of distributional change does not form two modes at any one year. That is, there is no *prima facie* evidence of the industrial missing middle phenomenon in Swaziland.

2.3.3.4 The "Missing Middle" is Missing in Swaziland

The analysis begins by looking at correlations between the average product of capital and firm-size for each year. A flat distribution of the average product of capital in relation to employment is observed from the data. This may be due to the high degree of capital irreversibility and adjustment cost of the plant and machinery component across establishments and over time. The observed empirical distribution suggests that; *ceteris paribus,* the marginal cost of capital is largely insensitive to firm size. Therefore, the evidence here shows that the return to capital is not scale dependent.

The investigation proceeds to plots nonparametric correlations between the average product of labour and firm size, and the relationship proves negative. This implies marginal costs of labour are declining with employment growth. A potential explanation here is that larger firms charge lower mark-ups due to scale economies in the domestic market compared to smaller firms. Another explanation relates to capital credit schemes provided by the State to small enterprises such as the Export Credit Guarantee Scheme that produce a mass of constrained firms which declines with firm size growth as access to capital credit diminishes. The more technical investigation of the average product of primary inputs using the dip-test established the absence of a bimodal distribution in both proxies of the marginal product of primary inputs.

2.4   Job Turnover and Measurement
2.4.1   Theoretical Measurement of Firm Turnover

This section describes and defines measurement concepts associated with job flows as applied to each size category of firms in the manufacturing sector in Swaziland. Drawing from Davis *et al*. (1996), the first index to consider for aggregating employment across firms is the weight, ω_{it}, presented as Eq. 1 in Table 2.6. It generalizes $\mathcal{E}_{jt}$ to refer to the set of plants in group j at time t or $t-1$, which includes drop-out and new firms. Here, a 'group' j may represent a region, a sector, plant-size category, or any other characteristics associated with firms.

Firms are also characterized in terms of entry-exit dynamics. Entrants are those establishments that have $N_{it-1} = 0$, $N_{it} > 0$ and $g_{it} = 2$, where N_{it} refers to the number of workers for firm i at time t. Firms that are exiting the industry are identified by $N_{it-1} > 0$, $N_{it} = 0$, and $g_{it} = -2$. At the firm-level, the firm growth index, g_{it}, refers to the rate of employment change between times t and $t-1$,

where the denominator is calculated as average employment. In the case of continuing and expanding plants, the characterization of firms requires that $N_{it} > N_{it-1} > 0$ and, $g_{it} > 0$, while those that are continuing, but contracting, are characterized by $N_{it} < N_{it-1}$ and $g_{it} < 0$. As a general principle, therefore, $g_{it} \in [-2, \ 2]$ and symmetric around zero, in contrast to the standard growth measure, G, that divides the lagged and current employment difference by the lagged employment level. The two measures are related by $G = \frac{2g}{2-g}$, see Davis *et al*. (1996a, p. 190).

On the basis of employment changes over time, job creation (JC) in Eq. 9 is defined as the combination of the sum of employment gains over all plants whose current employment level is greater than the previous period's level and the sum of employment gains from new plants. The new entrants are characterized by $N_{it-1} = 0$ and $N_{it} > 0$. Continuing and expanding firms include both young and mature firms identified by $N_{it} > N_{it-1}$. Conditional on survival, young plants are known to be smaller but grow faster through innovation and business entrepreneurship than the larger and more mature firms, see Neumark, Wall and Zhang (2011), Haltiwanger *et al*. (2013) and Decker *et al*. (2016).

Similarly, job destruction (JD) in Eq. 12 is associated with the contraction and closure of firms. It is essentially the sum of job losses over all firms whose current employment level is smaller than the previous period's level. Since exiting firms are defined as those with zero employment at time t, therefore JD captures both job decline at contracting establishments and job loss due to exiting establishments.

Net job reallocation (NR) or, simply put, net employment growth in Eq. 13 is the most preferred measure of job flows when interest is in the growth of the number of jobs. It is the difference between JC and JD. As illustrated in Davis *et al*. (1996a), this means that JC and JD decompose the net change of aggregate employment into a component related to expanding firms and a component connected with shrinking ones. For a given net employment growth, higher JC suggests the ease with which displaced workers and new labour market entrants find new jobs. Similarly, higher JD implies reduced job security for labour market participants.

Table 2.6: Measurement of Firm Turnover Indexes

INDEX	DEFINITION	EQUATION
Firm-Level Growth Rate	$g_{it} = (N_{it} - N_{it-1})/(1/2)(N_{it} + N_{it-1})$	(1)
Firm-Level Weight	$\omega_{it} = (N_{it} + N_{it-1})/\Sigma_{i\in\mathcal{E}_{jt}}(N_{it} + N_{it-1})$	(2)
Entry	$N_{it-1} = 0, \ N_{it} > 0,$ and $g_{it} = 2$	(3)
Exit	$N_{it-1} > 0, N_{it} = 0,$ and $g_{it} = -2.$	(4)
Continuing and Expanding	$N_{it} > N_{it-1} > 0;$ hence, $g_{it} > 0$	(5)
Continuing and Contracting	$N_{it} < N_{it-1};$ hence, $g_{it} < 0$	(6)
Gross Job Creation (JC_t):		
	$JC_{\text{Expanding}_t} = \Sigma_i \omega_{it}\max\{g_{it}, 0\};$	(7)
	$JC_{\text{Entry } t} = \Sigma_i \omega_{it}\max\{g_{it}, 0\}I\{g_{it} = 2\},$	(8)

	$JC_t = JC_{\text{Expanding}_t} + JC_{Entry_t}$	(9)
Gross Job Destruction (JD_t):		
	$JD_{\text{Contracting}_t} = \sum_i \omega_{it} \min\{-g_{it}, 0\}$	(10)
	$JD_{\text{Exit}_t} = \sum_i \omega_{it} \min\{-g_{it}, 0\} I\{g_{it} = -2\}$	(11)
	$JD_t = JD_{\text{Contracting}_t} + JD_{\text{Exit } t}$	(12)
Net Reallocation Rate (NR_t)	$NR_t = JC_t - JD_t$	(13)
Gross Reallocation Rate (JR_t)	$JR_t = JC_t + JD_t$	(14)
Excess Reallocation Rate (XR_t)	$XR_t = JR_t - |NR_t|$	(15)

(Gross) job reallocation rate (JR) in Eq. 14 is defined as the sum of JC and JD. In terms of Hijzen *et al*. (2010), the JR concept can alternatively be viewed as the highest number of worker movements required for adjustment to changes in job prospects across firms. This measure essentially counts workers both when they lose their jobs as a result of job destruction and also when they move to newly created jobs. That is, it consists of job gains from expanding plants and job losses from contracting ones. As observed by Davis *et al*. (1996a), other crucial characteristics of the labour market behaviour and performance may of course potentially vary with measures of job creation and destruction. For example, higher rates of JC and JD suggest greater heterogeneity in the reallocation of employment positions or jobs across firms. Such rates also mean larger numbers of workers are forced to reshuffle between jobs and also add to unemployment.

Finally, excess reallocation rate (XR) in Eq. 15 denotes the measure of the number of job changes in excess of those required to accommodate employment growth. The index XR can be calculated as the difference between JR and the absolute value of net employment change. This is an index of simultaneous JC and JD; see Davis and Haltiwanger (1999). The practical value of XR as a measure of job flow arises largely from its suitability for decomposition into two important components; namely, one that accounts for between-sector employment movements and another that accounts for excess JR within sectors, see Davis *et al*. (1996a: p.13) and Davis and Haltiwanger (1999: p.2717).

2.4.2   Industry, Employer Size and Job Flows: Empirical Findings

2.4.2.1 Job Creation, Destruction and Reallocation In Manufacturing

This section lays out the behaviour of job flows in the manufacturing sector in Swaziland relying on the size classification methodology developed by Davis and Haltiwanger (1992). The definitions presented in Table 2.6 on the basis of standard job flow nomenclature are used to measure industrial job flows in Swaziland.

The relationship between the gross job flows measures analysed in this chapter and size-weighted density distributions of firm growth observed in the literature (see, for example, Davis and Haltiwanger, 1992), is simple. It allows for the calculation of the sector-wide gross job creation by summing up employment gains arising from expanding incumbent firms and new entrants. The gross job destruction is calculated by summing up employment losses at contracting incumbent firms and

quitters. In Table 2.7, an analysis of job flows for the sector is conducted. We discuss the magnitudes together with the variation of job flows by firm size using the relations built up in section 4.1. The first and second columns report annual and cross-industry averages of job flows for plants employing at most 50 workers and plants with more than 50 workers. Columns 3 and 4 present annual and cross-industry averages of job destruction rates by firm-size category. Columns 5-10 respectively report similar information on net, gross and excess reallocation rates by firm size. A deeper analysis based on this table is conducted in the subsequent sections below.

Table 2.7: Rates of Job Creation, Destruction and Reallocation (1994-2003)

	Job Creation Rate		Job Destruction Rate		Net Job Reallocation Rate		Job Reallocation Rate		Excess Reallocation Rate	
Year	Firms≤50	Firms>50	Firms ≤50	Firms >50	Firms≤50	Firms>50	Firms ≤50	Firms >50	Firms ≤50	Firms >50
	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10
1995	0.0076	0.0203	−0.0908	−0.1876	−0.0832	−0.1673	0.0984	0.2080	0.0152	0.0407
1996	0.0178	0.0202	−0.0161	−0.0330	0.0017	−0.0128	0.0338	0.0532	0.0321	0.0404
1997	0.0097	0.0205	−0.0315	−0.1101	−0.0219	−0.0895	0.0412	0.1306	0.0193	0.0411
1998	0.0066	0.0202	−0.0169	−0.0117	−0.0104	0.0085	0.0235	0.0318	0.0131	0.0233
1999	0.0119	0.0353	−0.0127	−0.2407	−0.0008	−0.2055	0.0246	0.2760	0.0238	0.0705
2000	0.0104	0.0154	−0.0220	−0.0713	−0.0116	−0.0559	0.0325	0.0867	0.0209	0.0308
2001	0.0044	0.0205	−0.0112	−0.0359	−0.0068	−0.0155	0.0156	0.0564	0.0088	0.0409
2002	0.0065	0.0214	−0.0506	−0.1562	−0.0441	−0.1347	0.0571	0.1776	0.0130	0.0429
2003	0.0092	0.0534	−0.0714	−0.2127	−0.0621	−0.1593	0.0806	0.2661	0.0185	0.1068
Industry	**Firms≤50**	**Firms>50**	**Firms ≤50**	**Firms >50**	**Firms≤50**	**Firms>50**	**Firms ≤50**	**Firms >50**	**Firms ≤50**	**Firms >50**
Food (15)	0.0004	0.0087	−0.0012	−0.0232	−0.0012	−0.0145	0.0016	0.0319	0.0004	0.0174
Textile (17)	0.0021	0.0386	−0.0077	−0.2579	−0.0077	−0.2193	0.0098	0.2965	0.0021	0.0772
Apparel (18)	0.0273	0.0388	−0.0739	−0.0605	−0.0739	−0.0217	0.1012	0.0993	0.0273	0.0776
Wood (20)	0.0039	0.0152	−0.0173	−0.1654	−0.0173	−0.1502	0.0212	0.1806	0.0039	0.0304
Pulp & Paper (21)	.	0.0127	−0.0011	−0.2014	−0.0011	−0.1887	0.0011	0.2141	0.0000	0.0254
Printing & Publishing (22)	0.0089	0.0518	−0.0256	−0.0302	−0.0256	0.0216	0.0345	0.0820	0.0089	0.0604
Chemicals (24)	0.0051	0.0307	−0.0169	−0.2482	−0.0169	−0.2175	0.0220	0.2789	0.0051	0.0614
Rubber (25)	0.0315	0.0552	−0.0902	−0.0110	−0.0902	0.0442	0.1217	0.0662	0.0315	0.0220
Non-Metallic Mineral (26)	0.0124	0.0132	−0.0611	−0.1023	−0.0611	−0.0891	0.0735	0.1155	0.0124	0.0264
Basic Metals (27)	0.0923	.	−0.3508	−0.8185	−0.3508	−0.8185	0.4431	0.8185	0.0923	0.0000
Fabricated Metal (28)	0.0077	0.0272	−0.0286	−0.1035	−0.0286	−0.0763	0.0363	0.1307	0.0077	0.0544
Furniture (29)	0.0090	0.0698	−0.0516	−0.4514	−0.0516	−0.3816	0.0606	0.5212	0.0090	0.1396
Other Manufacturing (36)	0.0029	0.0280	−0.0107	−0.2147	−0.0107	−0.1867	0.0136	0.2427	0.0029	0.0560
TOTAL	**0.2035**	**0.3899**	**−0.7367**	**−2.6882**	**−0.7367**	**−2.2983**	**0.9402**	**3.0781**	**0.2035**	**0.6482**
MEAN	**0.0170**	**0.0325**	**−0.0567**	**−0.2068**	**−0.0567**	**−0.1768**	**0.0723**	**0.2368**	**0.0157**	**0.0499**
MEDIAN	**0.0083**	**0.0294**	**−0.0256**	**−0.1654**	**−0.0256**	**−0.1502**	**0.0345**	**0.1806**	**0.0077**	**0.0544**
STD DEV	**0.0256**	**0.0192**	**0.0929**	**0.2211**	**0.0929**	**0.2263**	**0.1177**	**0.2176**	**0.0250**	**0.0362**

Source: Author's calculations from data compiled by the CSO

A. Magnitude and Time Variation of Job Flows by Firm Size

The analysis begins with time-series patterns of job flows in the manufacturing sector. Using columns 1-4 in Table 2.7, Figure 2.5 plots the sum of employment gains from new entrants and expanding incumbents as well as the sum of employment losses from plants exiting and contracting. The simultaneity of job creation, JC_t, and destruction, JD_t, features throughout the period of analysis for each employer size category. However, JD_t significantly dominates JC_t in all plant sizes. That is, regardless of size, firms destroyed more jobs than they created. The high rates of average yearly job flows represent a persistent churning of job opportunities that characterizes the Swazi manufacturing labour market, cf. Davis and Haltiwanger (1999, Table 2) and Kerr *et al*. (2013).

Figure 2.5: Patterns of Job Creation and Destruction in Manufacturing (1994-2003)
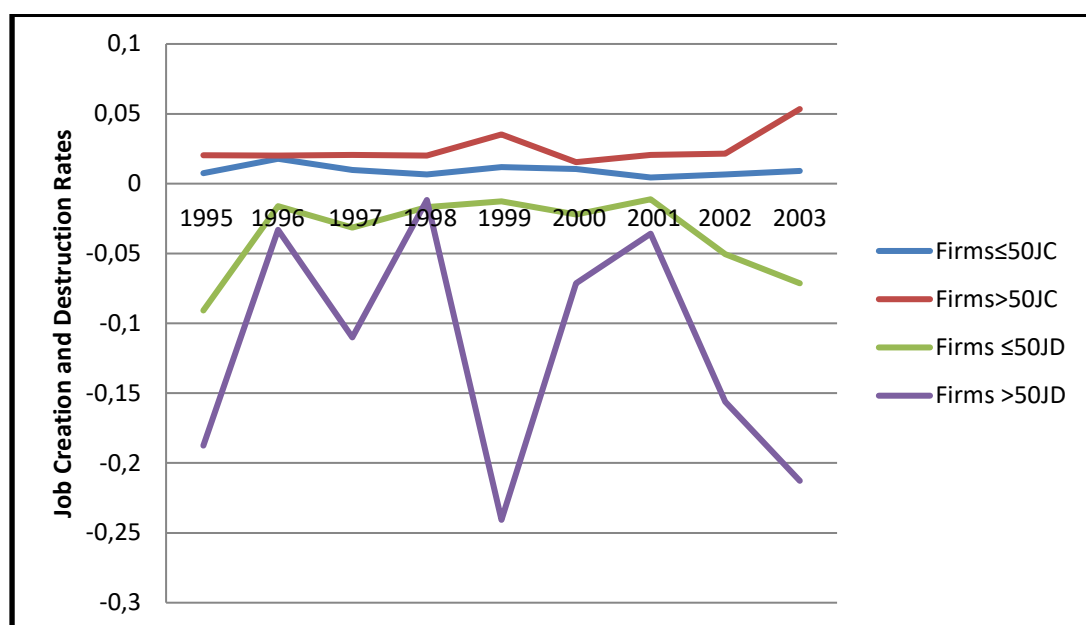
Looking at time series patterns in more detail, we find that the JC_t ability of small firms through expansion and turnover remained constrained for the entire period. Instead, small firms destroyed more jobs than they created. The contraction and exit of small firms peaked in 1995 and 2002/2003. Similarly, while large establishments portray no marked swings in JC_t, the average pattern shows positive shocks in 1999 and 2003. In contrast, there is rather significant activity concerning the contraction and exit of firms employing more than 50 workers. The JD_t process for large plants experienced during the reform period was pronounced in 1995, 1999 and 2003, and this coincides with specific events in the Customs Union; namely, the firm response to the South African political dispensation of 1994, mergers and acquisitions in the Pulp and Wood industry in 1998/1999 involving large firms, and the build-up to the end of the Multi-Fibre Agreement in January 2005.

A couple of remarks are helpful in thinking about the trends portrayed by job creation and destruction in the manufacturing sector in Swaziland. Large firms create and destroy more jobs than small establishments. In particular, Table 2.5 shows that small plants create only 0.9 percent jobs yearly while large plants create 3.5 times more new jobs every year. Job destruction by large firms shows sharp volatility throughout the period of analysis and it also exceeds the job destruction by small plants. Although small firms destroy an annual average of 3.59 percent jobs; large firms destroy approximately as much as 1 in 9 (or 11.77 percent) manufacturing jobs every year.[17] Thus, the manufacturing sector has been destroying more jobs than creating them as in Kerr *et al*. (2013) for South Africa's case.

The large-scale job reallocation activity observed in the manufacturing sector reveals a sense in which employment opportunities involving large plants change locations. This suggests consideration of a measure of simultaneity in the occurrence of job creation and destruction by firm size. By definition, excess reallocation is the gross job reallocation less the minimum amount required to accommodate the net change in manufacturing employment (see Davis *et al*., 1996). Figure 2.6 plots excess reallocation for both firm-size categories.

Figure 2.6: Patterns of Excess Job Reallocation in Manufacturing (1994-2003)
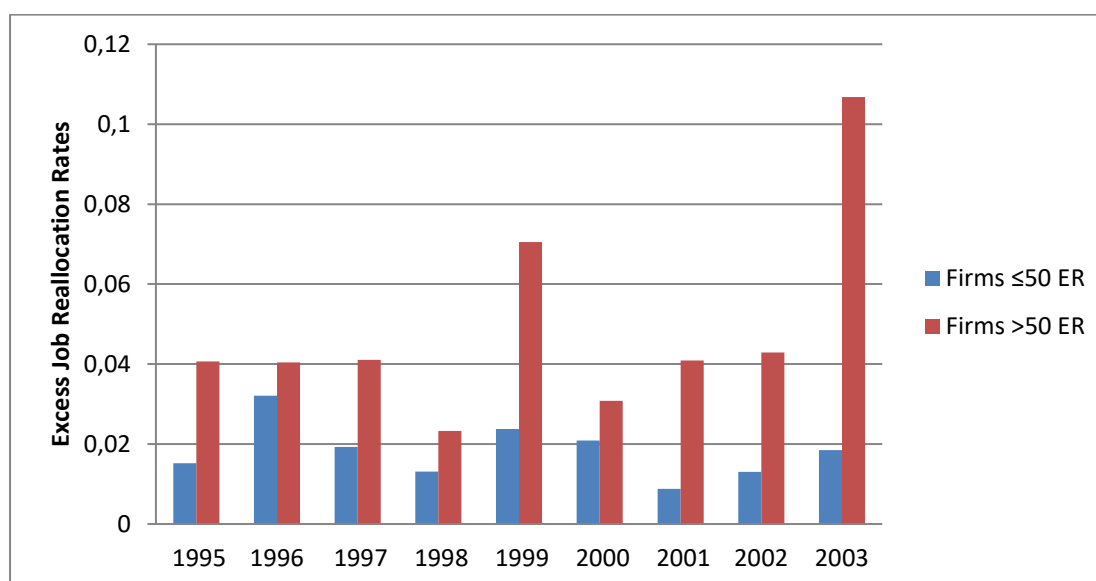
The relative dynamism in the large firm-size category finds full expression in excess job reallocation that dominates small establishments.

---

[17] The U.S. manufacturing sector created 1 in 10 jobs and also destroyed 1 in 10 jobs every year during the period 1973-1993 (Davis & Haltiwanger, 1999).

B. Cross Industry Variation of Job Flows by Firm Size

The cross-sectional presentation of job flows in the manufacturing sector allows us to assess their behavioural patterns across industries. Figure 2.7 reports the magnitude and variation of gross job flows by employer size and two-digit ISIC industry. The annual job flows exhibit higher average and median volatility for large plants. A notable result concerns the identification of industries that destroyed most jobs in the reference period. Focusing on industries generating more than 10 percent gross job destruction, we count nine out of 13 industries as high job destroyers by large plants and count 1 out of 13 industries as a high job destroyer by small firms.

Figure 2.7: Patterns of Job Creation and Destruction Rates in Manufacturing by Industry
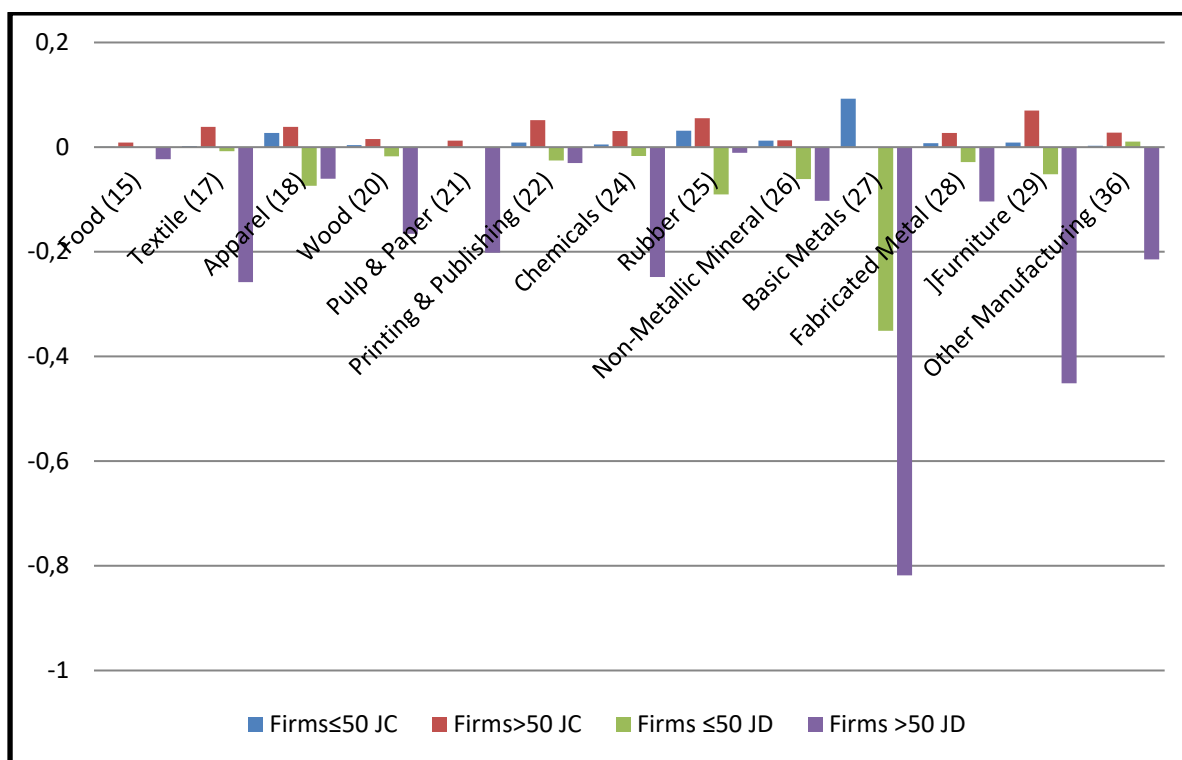
Job creation and destruction is also the basis for calculating the quantity that occurs beyond the measure that is required to bring about net sectoral contraction and expansion, that is to say excess job reallocation, Davis *et al*. (1996, p. 38). Figure 2.8 plots excess job reallocation by two-digit ISIC industries and it ranges from zero percent in the Pulp and Paper (21) to 3.15 percent in the Rubber (25) industries for small firms. This measure ranges from zero percent in the Basic Metals (27) to 13.96 percent in the Furniture (29) industries for large firms. The behaviour of small enterprises in job creation and destruction is in sharp contrast to that of larger plants. On the one hand, the uniformly low rates of excess reallocation for small plants suggest that every two-digit industry displays limited heterogeneity in the direction of employment growth. On the other hand, the highly dispersed excess

reallocation for larger firms indicates that a sizeable number of industries exhibit considerable heterogeneity in the employment growth. The interpretation for these results based on this industry classification scheme for large enterprises is that firm-level demand contains a substantial degree of uncertainty that is idiosyncratic to the individual firm [cf. Davis *et al.* (1996) for the U.S. case].

Figure 2.8: Patterns of Excess Job Reallocation in Manufacturing by Industry
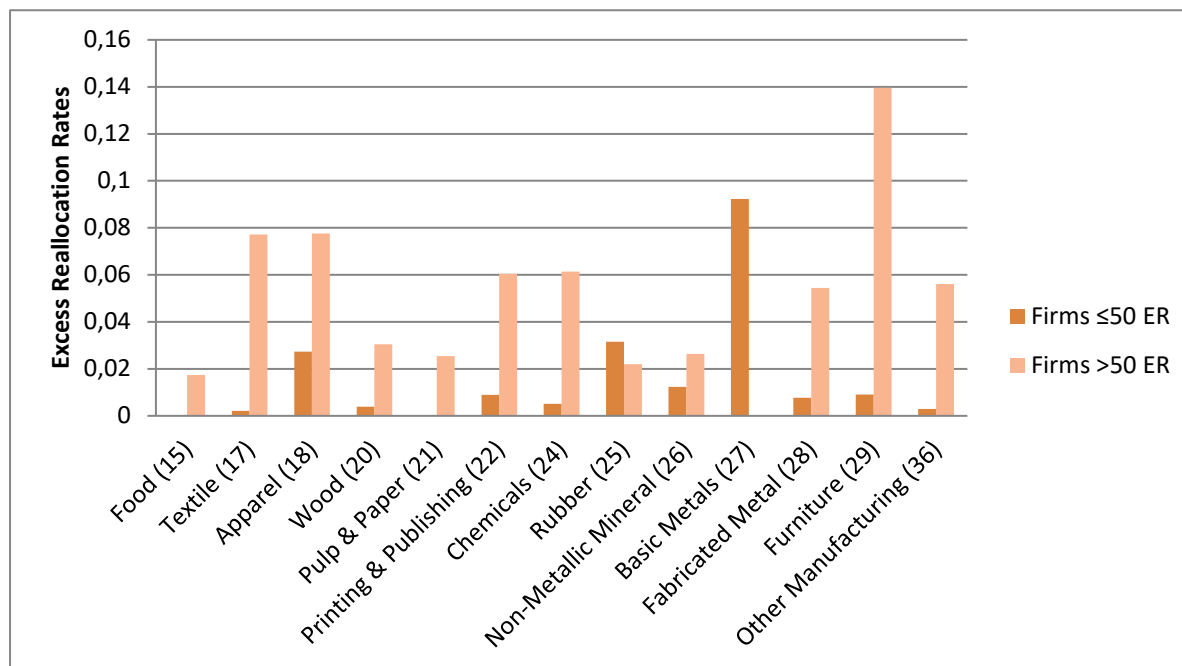
The rapid pace in the job reallocation in industries allows inference that a large proportion of gross job flows is induced by within-sector reallocation activity instead of by between-sector employment shifts.[18] Such results provide limited support for the view that high job reallocation rates are essentially caused by sectoral or economy-wide shocks.[19]

Evidence on job flows in developing countries is gaining momentum due to the increasing, though slow-paced, availability of micro-level data. In the African context, Shiferaw and Bedi (2009) examine gross job flows in the Ethiopian manufacturing firms, and find job creation and destruction patterns similar to those in developed and emerging markets. These results show higher job reallocation rates in industries dominated by young plants and start-ups. The same results obtain in other developing countries such as Cote d'Ivoire; see Klapper and Richmond (2011), who found that new entrants contribute disproportionately to gross job flows. Furthermore, both Klapper and Richmond (2011) and Shiferaw and Bedi (2009) find small and young firms more dominant in job

---

[18] Davis and Haltiwanger (1999) decompose the excess reallocation rate into two parts; namely, the within-sector reallocation and between-sector components. Empirical estimates of the decomposition overwhelmingly supports the within-sector reallocation hypothesis, see Hamermesh, Hassink and Ours (1996, Table 2) for Netherlands, and Doms, Dunne and Troske (1997, Table 5) for U.S.

[19] Cf. Acemoglu *et al.* (2012) and Gabaix (2011).

creation than large ones. Kerr, Wittenberg and Arrow (2013) examine the job creation and destruction patterns in the South African economy, and in the manufacturing sector in isolation. The incidence of job destruction was higher than that of job creation in 15 out of 24 waves; hence, the manufacturing sector was shedding employment rather than creating more jobs. Table 2.9 presents comparisons involving other countries. The manufacturing sector in Swaziland created only 2.5 percent new jobs while destroying 13.72 percent, consequently producing the worst net employment growth of -11.2 percent in the group. Although Swaziland destroyed the same percentage of jobs as the United Kingdom, the latter created and reallocated four times more jobs during period of global economic crisis than Swaziland during a period of trade liberalization.

Table 2.9: Comparative Job Flow Aggregates in the Manufacturing Sector across Countries

Country	Period	Business Unit	Job Creation	Job Destruction	Net Growth	Job Reallocation	Source
U.S.	1973-1993	Establishment	8.8	10.2	−1.3	19.0	Davis and Haltiwanger (1999)
U,K.	1997-2008	Enterprise	10.0	13.7	−3.7	23.8	Hijzen *et al.* (2010)
RSA	Wave 5-28	Enterprise	8.9	9.8	−0.9	18.7	Kerr *et al.* (2013)
Ethiopia	1997-2007	Firm	17.3	10.3	6.7	27.6	Shiferaw and Bedi (2009)
Swaziland	**1994-2003**	**Firm**	**2.5**	**13.7**	**−11.2**	**16.2**	**Author's Calculation**
Poland	1994-1997	Firm	3.3	5.0	−1.7	8.3	Faggio and Konings (2001)
Estonia	1994-1997	Firm	5.0	7.9	−2.9	12.9	Faggio and Konings (2001)
Slovenia	1994-1997	Firm	3.4	4.8	−1.4	8.2	Faggio and Konings (2001)
Bulgaria	1994-1997	Firm	2.4	5.6	−3.2	8.0	Faggio and Konings (2001)
Romania	1994-1997	Firm	3.0	8.1	−5.1	11.2	Faggio and Konings (2001)

In the South African manufacturing sector, Kerr *et al*. (2013) do not distinguish their industrial analysis according to job flow performance by small and large plants. However, they find that firms created nearly 9 percent jobs and destroyed 10 percent every year. Looking at the two-digit SIC level, the results remained robust. Job creation was still dominated by job destruction, except in the Food and Beverages, and Electrical Machinery industries. The study concludes that there was a general decline of about 7 percent in manufacturing employment between 2006 and 2011.

The cross-sectional heterogeneity in the magnitude of establishment-level employment adjustment has received a variety of explanations in the literature. First, as noted by Davis and Haltiwanger (1999); sector-specific shocks with differential effects among industries, cohorts of firm birth and firm size categories are potential drivers of gross job flows. Although the bulk of existing evidence suggests that sectoral shocks account for a small portion of gross job flows, Konings, Lehmann and Schaffer (1996) find that state-owned enterprises in Poland destroyed more jobs since the beginning of the transition to a market economy and job creation was dominated by the private sector. Second, the within-sector magnitude of heterogeneity indicates that idiosyncratic influences dominate the determination of which firms destroy and create jobs, which establishments achieve productivity growth and which ones suffer productivity declines. One reason advanced for such patterns involves the tremendous uncertainty about development, adoption, distribution, marketing and regulation of

new products and production techniques; see Davis and Haltiwanger (1999). Uncertainty concerning demand for new products or the cost efficiency of alternative production technologies encourages plants to experiment. Third, another reason put forward as a determinant of the observed heterogeneity in job flows involves differences in entrepreneurial and managerial ability that produce differences in net employment and productivity growth among firms.

2.4.3   Job Turnover and Labour Productivity Growth

This section is a very concise presentation of issues that are developed comprehensively in the next chapter. It focuses on job turnover that has proven to play a significant role in recent studies of aggregate labour productivity growth (ALP), see Hijzen, Upward and Wright (2010) and Brown and Earle (2002). Thus, the manufacturing sector in Swaziland is investigated to determine the extent of inter-industry productivity differences and the differential impact of decomposed components on aggregate labour productivity growth. We have defined firm size in terms of the number of workers employed. The productivity index is measured as a ratio of real value added (VA_{it}) to employment (L_{it}) for each establishment. Changes in productivity can be decomposed into within-firm effects, between-firm effects, and entry and exit effects, see De Loecker and Konings (2006) for an application to post-Socialist economies. Although there are a variety of different decompositions for productivity available in the literature, our choice is Forster *et al.* (2001). Armed with VA_{it} and L_{it}, it is straightforward to calculate the ALP growth index. Thus, plant *i*'s labour productivity at time *t* is represented by $\boldsymbol{\varphi}_{it} = \frac{VA_{it}}{L_{it}}$, the sector average ALP $(\boldsymbol{\varphi}_t)$ at time *t* and can then be expressed as $\boldsymbol{\varphi}_t = \frac{\sum_i VA_{it}}{\sum_i L_{it}} = \frac{VA_t}{L_t}$ while the employment share of plant *i* at time *t* is $s_{it} = \frac{L_{it}}{L_t}$. Movements in $\boldsymbol{\varphi}_t$ may reflect changes in embodied and disembodied technology as well as changes in technical efficiency

$$\boldsymbol{\varphi}_{jt} = \sum_{i \in j} s_{it} \boldsymbol{\varphi}_{it},$$

where s_{it} denotes the share of plant *i* in industry *j* and $\boldsymbol{\varphi}_{it}$ is the establishment-level measure of productivity. The change in $\boldsymbol{\varphi}_{jt}$ is decomposed as follows

$$\Delta\boldsymbol{\varphi}_{jt} = \left(\overbrace{\sum_{i \in C_t} s_{it-1} \Delta\boldsymbol{\varphi}_{it}}^{Within} \right) + \left(\overbrace{\sum_{i \in C_t} \Delta s_{it} * (\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})}^{Between} \right) + \left(\overbrace{\sum_{i \in C_t} \Delta s_{it} * \Delta\boldsymbol{\varphi}_{it}}^{Covariance} \right)$$
$$+ \left(\overbrace{\sum_{i \in EN_t} s_{it} * (\boldsymbol{\varphi}_{it} - \boldsymbol{\varphi}_{t-1}) - \sum_{i \in EX_t} s_{it} * (\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})}^{Net\ Entry} \right)$$

In this equation, C_t represents the set of continuing firms existing in two consecutive periods, EN_t is the set of start-up firms that are absent at *t-1* but enter the market at *t* and EX_t denotes the set of firms

that exit the market. The first term in the expression represents the component of aggregate labour productivity caused by efficiency improvements within establishments to produce more value added without an equivalent increase in labour input. The second component measures the between-firm differences in employment shares weighted by productivity at *t-1*. The third term is the cross-term measuring the product of changes in employment and productivity shares. The Net Entry component measures the combined effect of firm entry and exit on ALP.

In order to calculate the different components of ALP, we adopt the approach proposed by Petrin and Levinsohn (2012) and define ALP as the change in aggregate final demand minus the change in aggregate expenditures on primary inputs, see Nishida *et al*. (2014). This allows us to aggregate all establishment-level growth components to aggregate final demand. Data handling for achieving this involves double deflation of the value-added series following Bruno (1978) and Nishida *et al*. (2014) by using the manufacturing value-added deflator. We then calculate the different terms in the ALP growth expression for each four-digit industry and report results in Table 2.9 aggregated over two-digit ISIC industries.

Table 2.8: ALP Growth in Swazi Manufacturing Based on Foster *et al.* (2001) Decomposition (1994–2003).

ISIC2	Value-Added Growth	Labour productivity growth (0)	Within (1)	Between (2)	Cross (3)	Net Entry (4)
Food (15)	47.64	10.12	-4.99	4.41	-5.13	15.83
Textile (17)	47.72	11.24	-5.37	3.85	-4.27	17.02
Apparel (18)	56,15	16.42	-7.09	4.08	-3.60	23.04
Wood (20)	63.99	22.50	-5.02	3.98	-1.92	25.46
Pulp & Paper (21)	31.76	2.72	-7.36	5.26	-7.32	12.14
Printing & Publishing (22)	63.41	20.43	-3.92	2.52	-1.45	23.28
Chemicals (24)	51.70	10.67	-4.56	2.89	-3.51	15.85
Rubber (25)	70.41	20.85	-3.05	1.65	-0.79	23.04
Non-Metallic Minerals (26)	40.68	5.48	-5.03	4.21	-5.72	12.02
Basic Metals (27)	35.95	3.16	-6.35	5.67	-6.28	10.12
Fabricated Metal (28)	67.70	21.04	-3.19	2.30	-1.25	23.17
Furniture (29)	70.61	24.83	-3.44	2.31	-0.63	26.60
Other Manufacturing (36)	61.95	19.11	-4.07	2.78	-2.06	22.46
Mean	**54.59**	**14.51**	**-4.88**	**3.53**	**-3.38**	**19.23**
Median	**56.15**	**16.42**	**-4.99**	**3.85**	**-3.51**	**22.46**
Standard Deviation	**13.13**	**7.68**	**1.40**	**1.22**	**2.23**	**5.61**

Notes: The "Labour productivity growth" column depicts the ALP growth with entry and exit, and the "Value-added growth" column represents the aggregate real value-added growth rate. The plant-level real value added is summed and annualized across plants. As in Nishida *et al*. (2014), numbers are percentage growth rates. We define labour productivity as the amount of real value added relative to unit labour. $\Delta\varphi_t$ is decomposed into four components: (1) within, (2) between, (3) cross, (4) Net Entry term for the Foster *et al*. (2001) procedure. We use employment share as share weight, and both "within" and "between" terms use the base-period share for the weights.

The first column in Table 2.8 presents real value-added growth over the 10-year period. In the second column, we estimate the aggregate labour productivity growth which is then decomposed into the (1)-(5) components. We particularly focus our attention on the terms that combine to produce the

observed ALP growth, but begin with the generalizations presented using measures of central tendency at the bottom of the table. The deterioration of ALP growth comes from the firm-level inefficiency that negatively works to reduce the overall performance in the sector. As a result, the productivity effect of "Within-Firm" dynamics was on a typical year as low as -4.88 percent during the period of analysis. Since this average measure is higher than the median, it suggests the productivity distribution is left heavy-tailed with volatility estimated at 1.4. The lowest firm-level productivity growth is found in the Pulp and Paper industry at -7.36 percent followed by the Apparel industry at -7.09 percent, and the highest is in the Rubber industry at -3.05 percent. This can be interpreted as a reflection of a high level of unskilled workers in manufacturing with limited ability to convert technology into real output. The result is also consistent with the findings by Bloom *et al.* (2014) and Tybout (2000) that there is a prevalence of plant-level inefficiency and limited structural transformation in developing country manufacturing sectors.

The labour reallocation from low to high productivity firms presented by the "Between-Firm" column is on average 3.53 percent and positively skewed with volatility estimated at 1.22 which signifies a right heavy-tailed distribution. The Basic Metals and the Pulp and Paper industries recorded the highest level of ALP growth at 5.67 percent and 5.26 percent, respectively. However, the entry-exit dynamics produce the highest average productivity contribution to ALP growth at the net entry of 19.23 percent. The higher average productivity growth exhibited by the entrants' component indicates that there are extreme positive outliers pulling the mean over time. The resource reallocation and net entry dynamics reported here support the claim by Gelb *et al.* (2014) of the existence of productivity enclaves in a sea of small and lower productivity firms in Sub-Saharan Africa. As a whole, the labour productivity growth in the manufacturing sector in Swaziland is driven by resource reallocation across firms and firm turnover. The pattern of innovation and technological advancement works to reduce aggregate labour productivity growth. The latter suggests a deterioration of skill in the labour input and in managerial efficiency in this sector.

The magnitude of productivity differentials within and across sectors and economies, together with potential drivers of these, has recently received extensive attention. In Foster, Haltiwanger and Syverson (2008), U.S. industries manufacturing homogenous products such as oak flooring and cement exhibit 100 percent productivity spreads. Cross-country comparisons in the developing world seem to show magnified productivity differentials, with Hsieh and Klenow (2009) documenting the ratio of the 90th to 10th percentile of total factor productivity at 5 percent for Indian and 4.9 percent for Chinese firms, see Bloom *et al.* (2013). In a comprehensive study of 732 medium-sized firms in the U.S., France, Germany and the U.K., Bloom, Lemos, Sadun, Scur and Van Reenen (2014) collected managerial practice details and report a strong association of these with firm-level productivity and other outcomes. Bloom, Eifert, Mahajan, McKenzie and Roberts (2013) conducted a field experiment on large Indian textile firms to investigate the role of managerial practices in productivity and found

that management change raises productivity by 17 percent. Management practices are estimated by Bloom *et al*. (2014) to account for 25 percent of cross-country and within-country productivity gaps.

2.5   Discussion of the Results

 As industrial demand conditions, technological developments and trade policy evolved in Swaziland since the early 1990s, the microeconomic characteristics of the manufacturing sector responded cautiously in several dimensions. At the descriptive level, and looking at the firm size dynamics; the share of small firms remained approximately $\frac{2}{3}$ rds of the entire population survey of firms throughout the ten-year period. The related growth patterns conveyed through transition probabilities revealed a high degree of timidity of firms in terms of crossing thresholds of growth between the two size categories. That is, 98.02 percent of small firms remained small in the entire period while 98.28 percent of large firms remained large. This is in sharp contrast to the conventional wisdom that this firm growth dynamic is principally scale-dependent: small firms grow faster than large firms, conditional on survival; see Rossi-Hansberg and Wright (2007) and Decker *et al*. (2016). Thus, large firms in the manufacturing sector in Swaziland are effectively born large instead of emerging from transformational and fast growing small firms.

Turning to the entry/exit dimension of industrial characteristics provides an insight into plant-level churning and survival rates during the trade reforms. On average, 9.73 percent of firms entered the sector every year while failed establishments accounted for 8.04 percent. Firm survival in the sector was on average *at least* 90.27 percent, depending on whether survival is defined in terms of firms active either at $t-1$ and t or at t and $t+1$. The variability of these rates is quite pronounced, reflecting significant changes in the market structure precipitated by new import competition driving inefficient firms out of business while competitive ones entered the industry. Some firms ceased production for reasons other than weak productivity, but instead disinvested resources because of attractive larger markets presented by the reintegration of South Africa back into the world economy.

As an export-led economy, industrialization in Swaziland is largely driven by the extent of preferential foreign market access for its commodity products, see Edwards *et al*. (2013). The key export commodities that receive guaranteed non-reciprocal access to foreign markets include sugar, textile and apparel clothing traded in the European Union and the US. Beef exports are destined for Norway under the SACU-EFTA Trade Agreement have increased while larger volumes are traded locally and in Mozambique. The timing and occurrence of trade in foreign markets translated into increased and growing employment in these narrowly defined industries. At the aggregate two-digit industry level, Food is persistently the top employer in manufacturing. If the 2003 component of the data set is discounted as an outlier, the industry experienced an average employment growth of 8.38

percent and volatility of 7.91. It is followed by the Textile and Clothing industries, which are sensitive to foreign market access.

The preponderance of these firm dynamics during trade reforms, together with the associated labour demand, shaped the industrial structure in Swaziland. If the sector operated under dense regulatory conditions that increased fixed costs for large firms or under imperfect financial markets inaccessible to small firms, the expected outcome of the distribution of the marginal product of inputs would be the production of the 'Missing Middle' phenomenon. However, because of the non-observability of marginal products of primary inputs, and if the *average* product is proportional to the *marginal* product of inputs, the *average* product of inputs was analysed graphically and statistically. Firstly, the cross-sectional firm size lognormal distributions covering the 10-year period shifted towards the left. Secondly, this distribution evolves overtime to a state where it initially declines rapidly after its modal level, then slows down as if to form a 'dip' before accelerating again. The pattern of distributional change does not form two modes in any one year. Using a statistical approach following Hsieh and Olken (2014), we rely on Fan's (1992) design-adaptive nonparametric regression to determine the correlation between firm size and the average product of primary inputs. We regress the average product of capital and labour on firm size defined in terms of the log of employment. The dip-test statistic for both proxies of the marginal product of inputs strongly rejects the presence of bimodality in the ten annual cross-sections of the relevant data set.

Taking the data to a more rigorous analysis of longitudinal and cross-sectional job flows and firm size dynamics helps in our understanding of the labour market behaviour in each firm size category. This is important because of the tremendous churning that occurs in the Swazi industrial sector. Hundreds of jobs are created and destroyed every year as firms expand and prosper or old ones contract and shut down or as new plants enter the market. One key finding is that the job creating prowess of small industrial producers fails in the Swazi manufacturing sector. Instead, small firms destroy more jobs through the exit dynamic than they engage in job creation, with significant job description episodes experienced in 1995 and 2002-2003. In contrast, large firms created more jobs relative to small producers, they also experienced significant job destruction. More specifically, nine out of the 13 industries experienced job destruction in excess of 10 percent. This microeconomic churning process altered the industrial structure in Swaziland that obtained at the beginning of trade liberalization. The Basic Metals industry started shrinking abruptly from 1996, the Textile and Clothing industries expanded considerably in 2001 and the Pulp and Paper industry experienced gradual contraction since 1996. Are these industrial changes a reflection of firm-level responses to import competition or resource consolidation by South African global firms to intensify participation in the international economy along multiple margins to increase shares of global trade as suggested by Bernard *et al*. (2016)?

The cross-sectional decomposition of aggregate labour productivity identifies industries with the propensity of firm-level contraction and exit as a result of inefficient performance, and expansion and growth induced by higher productivity. The longitudinal dimension of the decomposition, or 'within-effects', shows the Pulp and Paper as the worst performer than the average industry, and that all industries have negative productivity. On the other hand, the reallocation effect of resources from inefficient to efficient uses is dominant; that is, it is persistently positive. Looking at both columns simultaneously, it is apparent that there is correlation between 'within-effects' and 'between-effects' of incumbent firms' productivity growth. The highest resource reallocation comes from the most inefficient industry and vice versa. This suggests that it is the less innovative firms that shed workers to high productivity innovative plants. Furthermore, firm turnover is consistently positive across all industries. This means that under-performing firms closed down operations and shifted labour shares to the new entrants with higher productivity growth.

A reconciliation of the results from gross job flow analyses with turnover effects shows that job destruction is scale-dependent: large firms account for most of the industrial gross job destruction. This gross job destruction is induced by inefficient large firms driven out of business by the more productive large entrants. In the Furniture industry, employment declines by 137 percent due to the exit of large low productivity firms in 2000 while the Metal Industry lost 139 percent in employment in 1996 for the same reasons. The general impact of firm turnover on aggregate labour productivity remains pronounced than growth effects from other growth components while large firms account for most job destruction through the exit margin. Some industries experience gradual contraction while others appear hit by transitory shocks at employment levels. Two potential explanations exist for this: first, large establishments engage in mergers and acquisitions occurring in a given year like in the Pulp and Paper industry. Second, a decision to disinvest may be taken at time t but because of high adjustment costs due to capital irreversibility and high hiring costs, it may take a while to exit the market.

2.6   Conclusion

Industrial firm size dynamics are scale-independent in Swaziland. Even if small firms survive, they do not grow faster than large firms. Transition probabilities show 98.02 percent of firms born small remain small after 10 years and 98.28 percent of firms born large remain large after the same period. Since the establishment turnover analysis shows a positive net entry of firms, the observed industrial employment growth does not come from incumbent firms increasing workers but rather from entry of new firms.

The notion that the distribution of firm size in developing African countries is characterized by a bimodal distribution with a missing middle is investigated graphically and statistically in the industrial

sector. The annual firm-size lognormal distributions shifted towards the left demonstrating a general economic deterioration during the 10-year period. This distribution evolves overtime to a state where it initially declines rapidly after its modal level, then slows down as if to form a 'dip' before accelerating again. The pattern of distributional change does not form two modes at any one year. Using a statistical approach, we regress the average product of primary inputs on firm size. The dip-test statistic for both proxies of the marginal product of inputs strongly rejects the presence of a missing middle in the ten annual cross-sections of the data. Thus, the missing 'missing middle' found in Hsieh and Olken (2014) for the cases of India, Indonesia and Mexico is confirmed for the case of Swaziland.

In the study of job flows, the manufacturing sector produces results that are consistent with findings in other countries. The simultaneity of gross job creation and destruction features throughout the period of analysis for each establishment size category. However, job destruction significantly dominates job creation in all plant sizes. Looking at longitudinal patterns in detail, the industry experienced a systematic failure of the job creating prowess of small firms in Swaziland. Small firms destroyed more jobs than they created them and large firms created more jobs than small plants. That is, while small plants create an annual average of 0.9 percent jobs, large plants create 3.5 times more new jobs every year. Although small firms destroy an annual average of 3.59 percent jobs; large firms destroy approximately as much as 1 in 9 manufacturing jobs every year. Thus, the manufacturing sector generally destroyed more jobs than it created them.

Interesting results are produced when job turnover is linked to productivity growth. The productivity effect of "Within-Firm" productivity was on a typical year as low as −4.88 percent during the period of analysis. This means that productivity growth coming from incumbent firms had growth-reducing effects on the overall productivity growth. Labour reallocation from low to high productivity firms presented by the "Between-Firm" column is on average 3.53 percent. This means that larger firms dominate the process of input resource reallocation to more efficient larger firms. However, entry-exit dynamics produce the highest average productivity contribution to ALP growth at the net entry of 19.23 percent.

Figure A.2.1: Average Product of Capital by Employment (1994-2003)

Graphs by Year
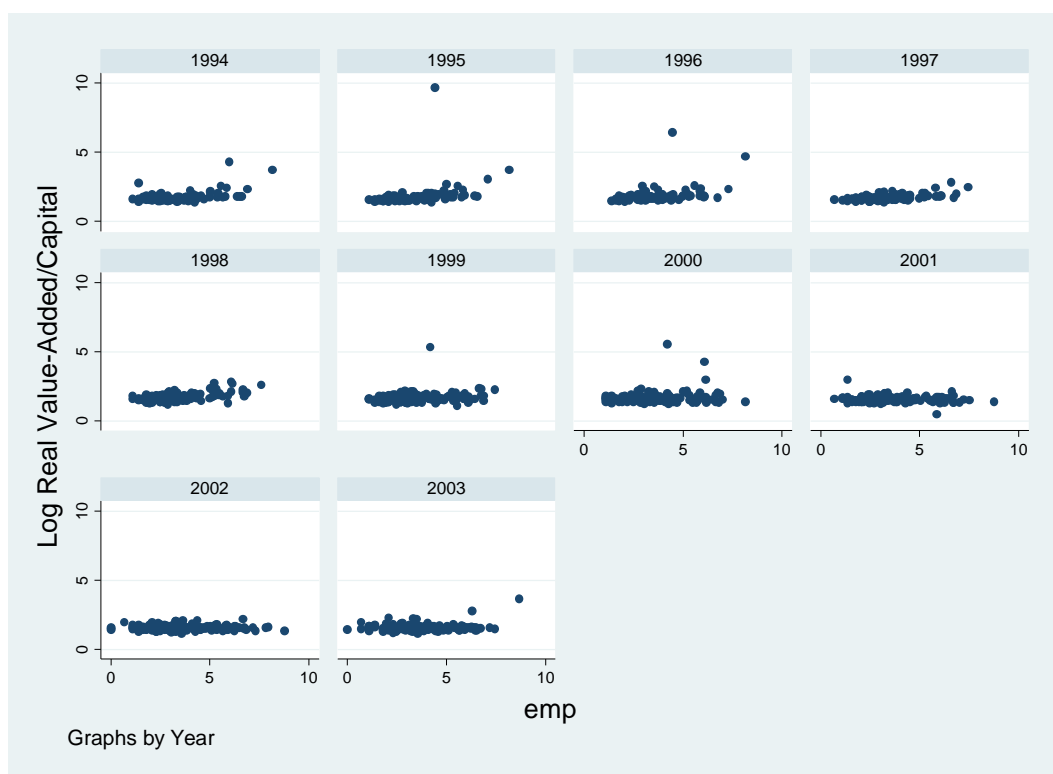
Table A.2.1: The Dip Test of the Average Product of Real Capital Stock

Year	Firms	Dip	*p*-Value	Low	High	Mean
1994	79	0.05	0.12	1.46	1.62	1.59
1995	85	0.03	0.44	1.53	1.64	1.59
1996	96	0.03	0.69	1.62	1.70	1.66
1997	99	0.03	0.44	1.66	1.86	1.76
1998	118	0.02	0.92	1.76	1.85	1.85
1999	139	0.03	0.35	1.47	1.55	1.51
2000	152	0.01	0.98	1.45	1.46	1.46
2001	165	0.02	0.80	1.48	1.49	1.49
2002	179	0.01	0.97	1.49	1.56	1.53
2003	151	0.02	0.86	1.49	1.55	1.52

49

CHAPTER 3: Does Technical Efficiency Dominate Resource Reallocation in Aggregate Productivity Growth?

3.1   Introduction

Recent research spurred by the increasing availability of longitudinal plant-level data, links microeconomic dynamics to aggregate outcomes. One area of focus for this research is the identification of establishment-level drivers and relative dominance of sources of aggregate productivity growth. A robust finding is that structural change effects of resource reallocation across plants are subordinate to within-plant productivity arising from learning-by-doing and learning-by-watching. For example, in nine of the 25 countries studied by Bartelsman, Haltiwanger and Scarpetta (2004) (or BHS) and Paǵes, Pierre and Scarpetta (2008) (or PPS), resource reallocation between plants was negative and weakly positive in only four countries. Similarly, in the analysis of job creation and productivity growth for the Slovenian manufacturing sector, De Loecker and Konings (2006) find dominance of technical efficiency over the reallocation of market-share of labour from low- to high-productivity incumbents as well as over firm turnover in driving aggregate productivity. In a comprehensive survey of the literature, Isaksson (2010) confirms for several countries at different stages of development that within-firm effects contribute more than inter-sectoral reallocation effects to aggregate labour productivity growth.

Another strand of the literature using enterprise-level micro-data also finds overwhelming evidence that within-industry reallocation of resources shape changes in industry aggregates; see Foster *et al*. (2008). This churning process and its effects on aggregate productivity have received special theoretical and empirical attention. As observed by Foster *et al*. (2008), models of selection mechanisms depict industries as assortments of producers characterized by heterogeneous productivity which link a firm's productivity level to its performance and survival in the industry. Key contributions in this area include Jovanovic (1982), Ericson and Pakes (1995), Melitz (2003), and Asplund and Nocke (2006). The main mechanism that causes change in these models is the reallocation of market-shares from either inefficient to efficient incumbent producers or from entry and exit of firms. Low-productivity establishments are less likely to survive and prosper relative to high-productivity incumbents which create selection-driven increases in industry productivity (Foster *et al*., 2008).

The common approach used to generate these results is largely based on the work of Baily, Hulton and Campbell (1992) and its derivatives such as Foster *et al.* (2001), Griliches and Regev (1995) and Olley and Pakes (1996). The Baily *et al.* (1992) method defines industry productivity growth as resource-share weighted changes in the distribution of the Solow-type technical efficiency (Solow, 1957). It derives its foundations from the decomposition and aggregation of plant-level residuals into

productivity growth components. The sources of this growth include changes in a plant's continuous innovation and adaptation to technological advances in the sense of learning-by-doing/watching as in Jovanovic (1982) and Pakes and Ericson (1998), movement in resource-share changes from low- to high-activity plants and turnover of firms. One question this method seeks to answer relates to the height of barriers to input reallocation in an economy, as in Bartelsman *et al.* (2004) and Paǵes *et al.* (2008).

The Petrin and Levinsohn (2012) method presents an alternative framework which introduces an environment with imperfect competition that creates a wedge in the marginal product−reward mix of inputs. It also creates a friction that induces heterogeneity in production technology and productivity levels, entry and exit of goods, costs of adjusting outputs and inputs, sunk and fixed costs, and markup-pricing. This is consistent with the recent work by Hsieh and Klenow (2009) and Petrin and Sivadasan (2013), who find significant heterogeneity between inputs' marginal products across establishments suggesting the presence of prohibitive distortions in input reallocation. Restuccia and Rogerson (2008) also calibrate a growth model with establishment-level heterogeneity arising from idiosyncratic policies and regulations, and institutional behaviour. This allows them to analyse the distortionary effects of such idiosyncrasies on the reallocation of resources across producers. Policies creating price heterogeneity among producers are found to reduce output and aggregate productivity by a range of 30 to 50 percent (see Restuccia and Rogerson, 2008).

The proposition by Petrin and Levinsohn (2012) has been applied by Nishida *et al*. (2014) to Chile, Colombia, and Slovenia; Ho, Huynh, Jacho-Ch´avez and Cubas (2014) to Ecuador, Petrin *et al.* (2011) to the U.S., and Kwon, Narita and Narita (2009) to Japan. This measurement approach defines aggregate productivity growth (hereafter referred to as APG) "as the change in aggregate final demand *minus* the change in aggregate expenditure on capital and labour" in the presence of imperfect competition and other distortions or frictions. Crucially, the APG decomposition has a term per establishment linked to technical efficiency and one for each primary input at each plant.[20] The term associated with either capital or labour is a function of the wedge between the value of the marginal product (VMP) and the relevant input price.

The purpose of this chapter is two-fold. *First*, it seeks to compare the individual drivers of aggregate labour productivity for the Swazi manufacturing sector with similar drivers for other countries. This exercise has never been done before for a Southern African country using a relatively long panel dataset compiled by a state agency.[21] *Second*, it estimates the components of industry productivity over time using both the Baily *et al.* (1992)/Foster *et al.* (2001) and Petrin and Levinsohn (2012)

---

[20] The phrase 'primary inputs' is used interchangeably with 'factor inputs'.

[21] Van Biesebroeck (2005) undertakes a similar analysis but has access only to RPED surveys, which have a short time dimension.

methods. In essence, the chapter examines the robustness of the overwhelming findings of the meta-analyses that productivity arising from learning-by-doing and learning-by-watching *dominates* productivity from market-share reallocation across incumbent firms and from net-entry of firms?[22,23] This question is examined across several dimensions using a rich and unique dataset for the manufacturing sector in a small developing African country- Swaziland.

This chapter makes three contributions to the literature. *First*, it applies the Baily *et al.* (1992)/Foster *et al.* (2001) approach to compare the drivers of industry productivity in Swazi manufacturing with similar growth drivers in Sub-Saharan economies, economies in transition and developed countries. *Second*, it uses the traditional approach and Petrin and Levinsohn (2012)/Nishida *et al*. (2014) to estimate ALP and APG over time. *Third*, it estimates the impact of confounding effects of plant turnover on the Baily *et al*. (1992) reallocation in the Swazi manufacturing data.

In the next section, we present an overview of the manufacturing sector in Swaziland for a period which coincides with trade liberalization and the political transition in South Africa. Section 3 undertakes descriptive analyses of key indicators and the behaviour of aggregate productivity for capital and labour. This is followed by a formal presentation of the Baily *et al.* (1992)/Foster *et al.* (2001) methodology for ALP decomposition in Section 4. Section 5 calculates ALP growth and its component drivers using the traditional method. In Section 6, we recast the Petrin and Levinsohn (2012)/Nishida *et al*. (2014) framework of APG decomposition, demonstrating how the Wooldridge (2009) modification of Levinsohn and Petrin (2003) is implemented in the dataset. Finally, we perform a direct estimation of APG and its component parts to determine the differential roles of technical efficiency and input reallocation on growth.

3.2   Overview of the Manufacturing Sector in Swaziland

The latter part of the 1980s was a period of unprecedented economic growth in the Swazi manufacturing sector. This was in response to economic sanctions on South Africa imposed by influential world economies (Edward *et al*. (2013)) and the relocation of some South African firms to neighbouring countries like Swaziland to circumvent these sanctions. The relocation decision enabled them to access foreign markets and/or to export intermediate inputs back to the home country. These foreign affiliates gained access to relatively cheap labour and material inputs in Swaziland, which reduced production costs. The domestic effect of this foreign presence in the sector came in the form of transfer of technical knowledge to local labour and to upstream suppliers. The resulting learning-

---

[22] Resource reallocation refers to reallocation of resources across incumbent firms and reallocation of resources in response to firm turnover, where firms/plants/establishments are used interchangeably.

[23] Levinsohn and Petrin (1999) refer to the learning-by-doing and learning-by-watching effects as the real productivity case.

by-doing increased both the efficiency of primary inputs and the quality of intermediate inputs from suppliers. Consequently, Hammouda, Karingi, Njuguna and Jallab (2010) found that Swaziland experienced 11.15 percent growth in real gross domestic product during the period 1985−1990 in which capital and total factor productivity accounted for 3.13 percent and 6.34 percent, respectively.

However, the period spanning the 1990s and 2000s was characterized by a marked deterioration in economic growth. This was due largely to the lifting of sanctions and re-integration of South Africa into the world economy (Hammouda *et al.,* 2010). In particular, trade liberalization that took place in the second half of the 1990s made South Africa appear as a more attractive investment destination. The response of South African multinational enterprises was to recall their foreign affiliates to improve their own scale economies, see Jonsson and Subramanian (2001). As international competition intensified, domestic industries that were characterized by oligopolistic markup pricing behaviour were forced to behave competitively. According to Jonsson and Subramanian (2001), the consequence of a freer market environment was the exit of some of the inefficient firms which, in turn, reallocated market shares to continuing ones and also to industry entrants. They also argue that, despite the presence of such import discipline mechanism, the limited domestic market size still enabled a portion of inefficient plants to survive and also allowed new low-productivity manufacturers to enter the market.

During this period the Swaziland Government responded with an attempt to address the issue of missing markets in the economy. One critical area for industrial policy intervention involved institutional reforms and infrastructure development to attract FDI, see Masuku and Dlamini (2009). As a result, the Swaziland Industrial Development Corporation (SIDC) was commissioned to design and implement a factory shell development programme to reduce sunk investment costs for producers, particularly in the textile and apparel industries. The Swaziland Investment Promotion Authority (SIPA) was also established in 1998 as a one-stop shop to serve mainly foreign investors. The objective of SIPA's existence was to market the country abroad as an investment destination and also to serve as an information desk when the foreign firm was ready to invest in Swaziland. In addition to these efforts to lure foreign investment, the state was also an active participant in the domestic economy. Direct state presence through Tibiyo TakaNgwane sought, *inter alia*, to increase formal sector employment and earn foreign exchange.[24] The presence of this state-owned enterprise is found in key sectors of the economy, and is perceived by the Federation of Swaziland Employers and Chamber of Commerce as having undesirable crowding-out effects on private investment, see Tibiyo TakaNgwane's Annual Report (2010).

---

[24] Tibiyo TakaNgwane is a state-owned enterprise whose purpose is to actively pursue commercially viable projects in all sectors of the economy (Tibiyo TakaNgwane Annual Report, 2012).

3.3 Descriptive Analysis of the Panel Data Series

3.3.1 Data Description and Summary Statistics

Although a detailed account of the source and structure of the dataset is presented in the overview chapter, the investigation of aggregate productivity growth requires a more direct description of relevant data series. Firm dynamics in the 1990s and early 2000s were driven by an average entry rate of 9.72 percent and exit rate of 8.03 percent per year. In the same period, the aggregate labour series oscillated around an average of 21 500 employees as shown in Table 3.1. In particular, aggregate labour changes exhibit relatively erratic patterns of weakening over the entire period. At the same time, the real value-added series in column four was largely static, except for a sharp drop in 1997.

Table 3.1: Summary Statistics

Year	No. of Firms	Employment	Total Amount in E' Million*		
			Real Value-Added	**Real Capital**	**Real Wages**
1994	100	17 260	1 241.28	2 221.8	2 267.2
1995	109	18 216	1 033.25	1 445.7	2 144.4
1996	117	17 837	1 132.92	1 085.7	2 271.0
1997	130	18 513	1 164.43	1 287.2	1 433.3
1998	150	20 296	1 087.23	2 928.2	1 605.6
1999	153	19 760	2 568.00	5 344.3	2 042.5
2000	164	19 036	2 291.59	5 477.6	2 705.8
2001	177	28 861	2 697.51	5 482.2	2 685.7
2002	188	32 219	2 143.96	6 879.5	2 830.7
2003	160	23 499	1 919.77	6 557.2	2 852.9
Mean	144.8	21 550	1 727.99	3 871.9	2 283.9

Note: * These figures were transformed using double-deflation of value-added, capital and the wage series as required by Bruno (1978) and applied by Nishida *et al.* (2014) for the case of Chile, Colombia and Slovenia.

The events that characterize the churning process of firms included the deepening pressure for higher wage increases by unions, and the resulting worker unrest necessitated restructuring of businesses through retrenchments.[25] Industrial action was however more visible in some sectors than in others. Moreover, the increase in aggregate capital was rather rapid from 1996 and levelled off somewhat in 1999. Since capital measurement is based on the plant, machinery and equipment (PME) series, which excludes the cost of repairs and replacement, its years of upward trend is a reflection of generally lumpy investment in fixed capital by a few large firms.[26]

---

[25] See the Central Bank of Swaziland Reports (1995-2003) on industrial unrests and IMF Staff Report (2000:13) on the need to review the Industrial Relations Act.

[26] The intermittence and lumpiness of capital projects as well as indivisibility contribute to the non-smoothness in the adjustment path of capital stock; see Nielsen and Schiantarelli (2003). Indivisibility ensures that investment occurs only in discrete increments.

3.3.2 Aggregate Input Productivity Movements

Aggregate input productivity changes in manufacturing during the trade liberalization period show a general decline as shown in Figure 3.1. The aggregate labour productivity index mimics aggregate labour input trends examined above. This suggests a high level of co-movement between value-added output and aggregate labour productivity. It is therefore not surprising to see a rapid decline in aggregate capital productivity from a point in time when the capital series begins an increase. Furthermore, the capital-labour ratio shows an increase after the first three years. This reflects a general increase in capital-intensity in production during the period under analysis without corresponding growth in real value added.

Figure 3.1: Output-Input and Capital-Labour Ratios by Year
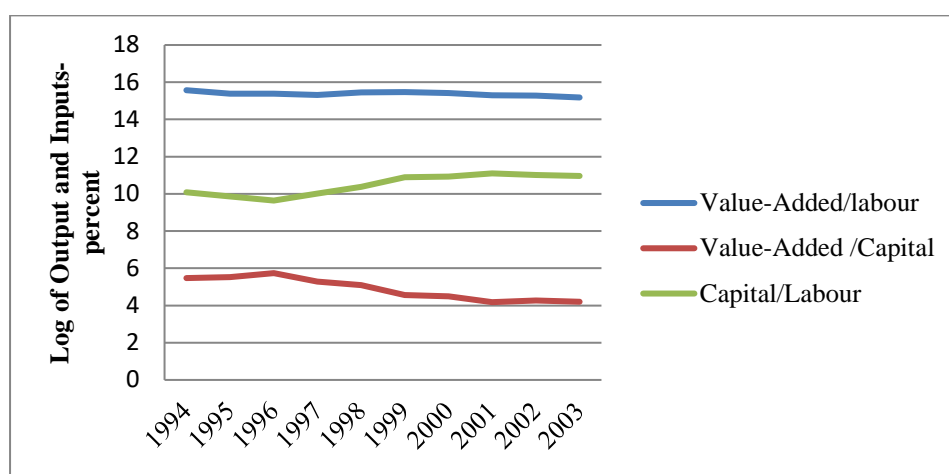
Figure 3.2: Output-Input and Capital-Labour Ratios by Industry (1994-2003)
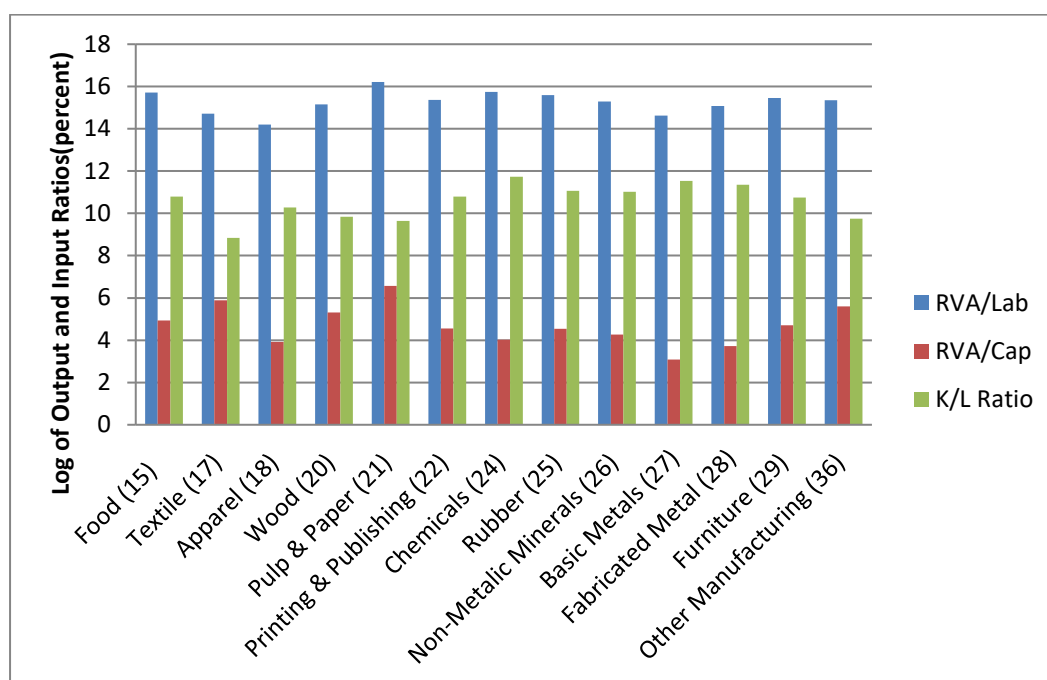
In general, the descriptive analysis is consistent with an explanation where a significant proportion of larger firms shed labour and keeps capital adjustment levels largely unchanged. This pattern of firm behaviour aligns with an economic environment which favours shifting most of the production by South African affiliates in Swaziland back to South Africa. Given that capital is mostly irreversible in nature, these firms could not recoup the fixed costs of capital but simply operated to cover their variable costs to remain in business.

This evidence sheds light on average patterns of aggregate factor input productivity across time and industry but cannot reveal much, if anything at all, about its cross-sectional distribution at a given point in time. Looking at aggregate labour productivity (ALP), Figure 3.2 shows a persistent shift of ALP towards the left with growing fat-tails in both directions. These patterns remain unaltered even when the value-added series is subject just to single deflation, except that the whole distribution moves more to the left, see Appendix A3.2. This is in sharp contrast to conventional wisdom, which holds that market reforms increase productivity within and across firms to drive aggregate growth.[27] Normally, trade liberalization has been shown to increase firms' incentives to invest in innovative technologies, and weak firms to lose market share to efficient ones, thereby boosting productivity, see Lileeva (2008).

Figure 3.3: ALP Distribution for Selected Years (1994, 1997, 2000, 2003)

---
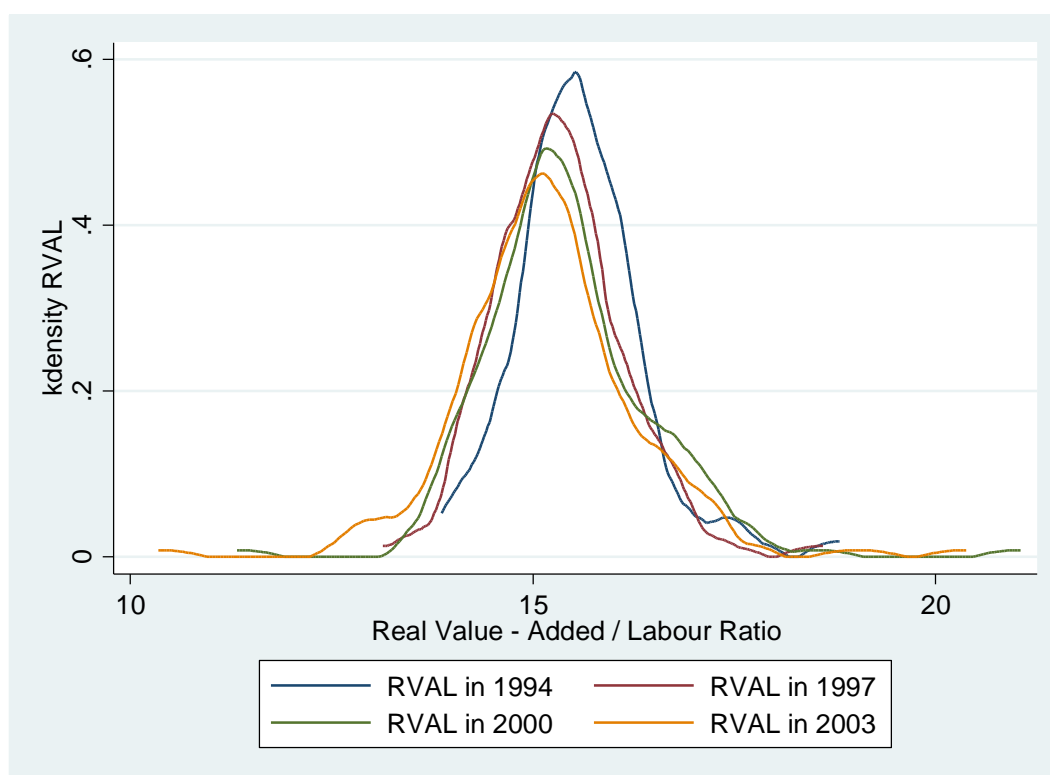
[27] See; for example, Lileeva (2008, Fig.1) for the case of Canada within NAFTA where the evolution of growth generated from the 'Between' and 'Within' terms continuously shift towards the right. Escribano and Stucchi (2014, Fig.1) find productivity improvement for Spanish manufacturing firms during a recession.

Note: ALP is measured as a ratio of double-deflated value added to aggregate employment in a year; see Appendix A3.2 for a single-deflated ratio of real value added to annual total employment.

In Table 3.2, we report patterns of productivity index movements by industry, and measure their central tendencies and dispersion. This allows us to document the relative performance of industries in relation to the chosen base year. Our first year of the sample period −1994− is normalized to one and the productivity index for the subsequent years is measured relative to this base year. On average, there is at best stagnation in 1998-1999 and at worst a loss of about 3 percent in productivity by 2003. This is contrary to De Loecker and Konings (2006) who use Olley and Pakes (1996) to find an average increase of 63 percent in the productivity index for Slovenia covering the period 1994-2000. The presence of heterogeneity is starkly reflected by a 2 percent growth in the 'Wearing Apparel' industry, while the 'Basic Metals' industry declines by 9 percent in the final year. Again, De Loecker and Konings (2006) found increases of 7 and 77 percent in the respective industries. However, the Pulp and Paper industry remains the dominant driver of ALP growth in the trade reform period in Swaziland.

Table 3.2: Evolution of the Average ALP by Industry (1994-2003)

Industry	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Food and Food Products	1.00	0.97	0.97	0.98	0.99	0.98	0.98	0.98	0.97	0.96
Textile	1.00	0.98	0.98	0.98	1.00	1.00	0.97	1.00	0.97	0.96
Wearing Apparel	1.00	0.99	1.03	1.01	1.03	1.01	1.01	0.88	1.00	1.02
Wood and Wood Products	1.00	0.97	0.96	0.96	0.99	0.98	1.00	0.97	0.98	0.95
Pulp and Paper Products	1.00	1.01	1.03	1.05	1.05	1.05	1.05	1.01	1.02	0.97
Printing, Publishing	1.00	0.99	0.99	0.99	0.99	0.99	1.00	1.00	0.99	0.99
Chemicals Products	1.00	0.99	1.00	1.01	0.98	0.99	1.00	1.01	1.00	0.98
Rubber and Plastic Products	1.00	0.99	0.99	0.97	0.99	0.98	0.97	0.98	0.96	0.94
Other non-metallic Minerals	1.00	0.99	0.96	0.94	0.96	0.97	0.98	0.96	0.98	0.97
Basic Metals	1.00	0.99	1.01	1.02	1.01	1.02	0.98	0.98	0.99	0.91
Fabricated Metal Products	1.00	0.99	1.01	0.98	0.99	0.99	0.97	0.97	0.98	0.99
Machinery and Equipment	1.00	0.99	0.99	0.99	0.99	0.99	0.97	0.99	1.00	0.99
Furniture	1.00	1.02	1.00	0.98	0.98	1.02	0.98	0.99	0.98	0.98
Sector Mean	**1.00**	**0.99**	**0.99**	**0.99**	**1.00**	**1.00**	**0.99**	**0.98**	**0.99**	**0.97**
Sector Median	**1.00**	**0.99**	**0.99**	**0.98**	**0.99**	**0.99**	**0.98**	**0.98**	**0.98**	**0.97**
Std Dev (σ_{ALP})	**0.00**	**0.01**	**0.02**	**0.03**	**0.02**	**0.02**	**0.02**	**0.03**	**0.02**	**0.03**

Source: Author's calculations.

It also seems natural to perform an analysis of ALP behavioural patterns at the tails of its distribution. For example, the 25th percentile in the ALP distribution shows more volatility than either the average situation or the upper 75th tail. That is, the standard deviation of the 25th percentile was $\sigma_{ALP} \in [1.05, 6.68]$ whereas the 75th percentile was characterized by $\sigma_{ALP} \in [0.11, 0.26]$ as shown in Appendices A3.3 and A3.4, respectively. This suggests that a firm in the 25th percentile ALP distribution was more sensitive to productivity shocks than either an average or a third-quartile firm. As a result of these industrial productivity swings, the bottom and 75th percentile firms experienced an ALP decline of 8 and 5 percent, respectively.

50

The emerging ALP trends and heterogeneity suggest the need for a deeper understanding of microeconomic causes and foundations for productivity growth, or in Swaziland's case stagnation and decline. It is therefore necessary to disentangle the roles of real productivity, intensive margins of share-shift effects, and extensive margins of turnover in productivity growth across industries. We achieve this in the next section by formally presenting a framework that outlines the relationship between resource shares and the productivity index in calculating each component of the ALP decomposition.

3.4 Measurement and Decomposition of Aggregate Labour Productivity

3.4.1   Definition and Measurement of ALP Growth

The quantity of labour (L_{it}) as a primary input in production at firm i is measured by the head-count of paid workers and working proprietors.[28] Nominal value-added output is measured as gross output *minus* intermediate inputs; that is, material and energy. Following Nishida *et al*. (2014) and Petrin *et al*. (2011), the quantity index of real value added (VA) is then constructed by using the double-deflation approach to nominal value added proposed by Bruno (1978) as

$$VA_{it} = \frac{P_{it}Q_{it}}{P_t^Q} - \frac{P_{iMt}M_{it}}{P_t^M} - \frac{P_{iEt}E_{it}}{P_t^E} \tag{1}$$

$$\cong \frac{P_{it}Q_{it} - P_{iMt}M_{it} - P_{iEt}E_{it}}{P_t}$$

where Q_{it}, M_{it} and E_{it} are nominal gross output and inputs of material and energy with their respective price indices. The double-deflation expression in the first line of Eq.1 represents the relevant price index for gross output and intermediate input quantities, see Petrin *et al*. (2014, Appendix 3) for Chile. The second line of Eq.1 presents the expression of a single-deflation method approximated with a common industry price deflator for both the output quantity and intermediate inputs, see Petrin *et al*. (2014, Appendix 3) for Colombia and Slovenia. The single-deflation approach is useful whenever intermediate deflators are not available.

Armed with information on VA_{it} and L_{it}, it is straightforward to calculate the ALP growth index. Thus, plant *i*'s labour productivity at time *t* is represented by $\boldsymbol{\varphi_{it}} = \frac{VA_{it}}{L_{it}}$ and aggregate labour

---

[28] The best measure of labour input according to OECD (2001) is hours worked. Although the legal length of a work-day is 8 hours and public holidays are known for the Swazi manufacturing sector, there is no information on worker absenteeism, variation in overtime, evolution of part-time work, sick leave and employee slack time due to ill-health. Furthermore, in the absence of the total number of hours worked that can be divided by the average annual number of hours actually worked in full-time jobs, the use of full-time equivalent employment is not feasible for the labour input definition contained in Doraszelski and Jaumandreu (2013, Appendix A) and OECD (2001).

productivity $(\boldsymbol{\varphi}_t)$ at time t can then be expressed as $\boldsymbol{\varphi}_t = \frac{\sum_i VA_{it}}{\sum_i L_{it}} = \frac{VA_t}{L_t}$ while the employment share of plant i at time t is $s_{it} = \frac{L_{it}}{L_t}$. Movements in $\boldsymbol{\varphi}_t$ may reflect changes in embodied and disembodied technology as well as changes in technical efficiency.[29] These changes may also reflect shifts in scale economies and degrees of capacity utilization. For the decomposition of aggregate labour productivity growth, $\Delta\boldsymbol{\varphi}_t$, the literature relies largely on the tradition of Baily *et al.* (1992)/Foster *et al.* (2001) in defining the effects of its sources. Specifically,

$\Delta\boldsymbol{\varphi}_t$ = Within Effects + Between Effects + Cross Effects + Net Entry Effects

$$= \Delta\boldsymbol{\varphi}_{WEt} + \Delta\boldsymbol{\varphi}_{BEt} + \Delta\boldsymbol{\varphi}_{Cross} + \Delta\boldsymbol{\varphi}_{Net-Entry}. \qquad (2)$$

Eq. 2 means that aggregate labour productivity growth, $\Delta\boldsymbol{\varphi}_t$, increases when firms use innovative production methods to produce more output through the 'Within-Firm' effects term $\Delta\boldsymbol{\varphi}_{WEt}$, holding factor inputs constant. The $\Delta\boldsymbol{\varphi}_t$ index can also increase when inefficient incumbent firms reallocate resources to more efficient ones through the term $\Delta\boldsymbol{\varphi}_{BEt}$. Haltiwanger (1997) adds a component that allows for the interaction between the change in resources and the change in ALP growth, and calls it the 'cross' or the 'covariance' term. The cross term increases when the changes in both components move in the same direction; that is, when the market share and ALP growth jointly increase and vice versa. Lastly, if new business methods including capital deepening that lead to improvements in industry productivity can only be adopted by new plants, then the net-entry term, $\Delta\boldsymbol{\varphi}_{Net-Entry}$, should dominate.

Motivated by PPS and BHS, Nishida *et al.* (2014) perform a theoretical and empirical analysis of ALP growth and APG using traditional methods and Petrin and Levinsohn (2012), respectively. We replicate Nishida *et al.* (2014) for the case of the manufacturing sector in Swaziland by decomposing ALP on the basis of Baily *et al.* (1992)/Foster *et al.* (2001) and APG using the marginal product of factor inputs.

3.4.2 The ALP Growth Decomposition Using the Baily *et al.* (1992) Method

The traditional method of $\Delta\boldsymbol{\varphi}_t$ decomposition is associated with the Baily *et al.* (1992) approach and its derivatives such as Griliches and Regev (1995), Foster *et al.* (2001) and Olley and Pakes (1996). In this context, $\Delta\boldsymbol{\varphi}_t$ is traditionally defined as input-share weighted changes in the distribution of plant-level technical efficiency, covariance and resource reallocation by incumbents and net entrants into the market. The Baily *et al.* (1992) decomposition additively isolates $\Delta\boldsymbol{\varphi}_t$ gains arising only from

---

[29] Embodied technology refers to advances in the design and quality of new vintages of capital goods and intermediate inputs, and disembodied technology refers to new blueprints, scientific results and new organizational techniques, see OECD (2001).

technical efficiency and resource reallocation. The Baily *et al.* (1992) (or BHC_{ALP}) procedure decomposes $\Delta\boldsymbol{\varphi}_t$ as

$$BHC_{ALP} = \left(\overbrace{\sum_{i\in C_t} s_{it-1}\Delta\boldsymbol{\varphi}_{it}}^{Within}\right) + \left(\overbrace{\sum_{i\in C_t} \Delta s_{it} * \boldsymbol{\varphi}_{it-1}}^{Between}\right) + \left(\overbrace{\sum_{i\in C_t} \Delta s_{it} * \Delta\boldsymbol{\varphi}_{it}}^{Covariance}\right) +$$

$$\left(\overbrace{\sum_{i\in EN_t} s_{it} * \boldsymbol{\varphi}_{it} - \sum_{i\in EX_t} s_{it-1} * \boldsymbol{\varphi}_{it-1}}^{Net\ Entry}\right) \tag{3}$$

where $\Delta\boldsymbol{\varphi}_{it} = \boldsymbol{\varphi}_{it} - \boldsymbol{\varphi}_{it-1}$ and $\Delta s_{it} = s_{it} - s_{it-1}$, and EN_t and EX_t represent firm entry and exit at time *t*, respectively. The different sources of $\Delta\boldsymbol{\varphi}_t$ are defined as

Within-plant effects: $\sum_{i\in C_t} s_{it-1}\Delta\boldsymbol{\varphi}_{it}$ is the sum of changes in plant-level labour productivity weighted by *t−1* base-period labour share for continuing plants. It measures a plant's gains in productivity induced by continuous improvement in production methods without an increase in its labour share. This growth component is referred to as real-productivity effects in Levinsohn and Petrin (1999).

Between-plant effects: $\sum_{i\in C_t} \Delta s_{it} * \boldsymbol{\varphi}_{it-1}$ in Baily *et al.* (1992) is the sum of changes in plant-level employment shares multiplied by the *t−1* labour productivity for continuing plants. This growth effect measures the extent of labour share reshuffling across plants where the labour input is reallocated to more efficient plants. This term is also viewed as 'clean' because it holds real productivity constant; see Nishida *et al.* (2014).

Covariance effects: $\sum_{i\in C_t} \Delta s_{it} * \Delta\boldsymbol{\varphi}_{it}$ is the sum of plant-level contemporaneous changes in the labour share and labour productivity. As Nishida *et al.* (2014) point out, this term increases when plants with increasing labour productivity are also plants with increasing labour shares.

Net-entry effects: An entering plant is identified when it first appears at time *t*, and an exiting plant is identified when it last appeared at time *t−1*. Thus, for $\sum_{i\in EN_t} s_{it} * \boldsymbol{\varphi}_{it} - \sum_{i\in EX_t} s_{it-1} * \boldsymbol{\varphi}_{it-1}$, where $\boldsymbol{\varphi}_{it}$ enters the equation as raw data for firm i at time t, positive contributions to ALP growth arise from the entry of high productivity firms and exit of inefficient ones. Net-entry effects therefore refer to the difference between productivity growth contributions by entering and exiting plants.

In the BHC_{ALP} formulation of resource movement between plants in Eq. 2, as Forster *et al.* (2001) and Nishida *et al.* (2014) point out, even if all plants have the same level of productivity for both the beginning and end period, the between component and net-entry component will in general be nonzero. Moreover, previous studies such as Syverson (2004) have estimated high dispersion in measured productivity, which translates to large and volatile (Baily *et al,* 1992) 'Between' effects.

The standard remedy for this is to 'normalize' each industry's 'Between' and 'Within' terms by the industry's ALP and use the industry's revenue shares as weights to aggregate across industries, see Petrin and Levinsohn (2012). As in Petrin and Levinsohn (2012) and Nishida *et al*. (2014), no normalization is carried out here in order to avoid losing the potential link between the actual ALP and BHC_{ALP}, although the nature of such a link prior to normalization is unknown[30].

3.4.3   The ALP Growth Decomposition Using the Foster *et al.* (2001) Method

The decomposition of ALP using Foster *et al.* (2001) (or FHK_{ALP}) is given as

$$FHK_{ALP} = \left(\overbrace{\sum_{i \in C_t} s_{it-1} \Delta \boldsymbol{\varphi}_{it}}^{Within} \right) + \left(\overbrace{\sum_{i \in C_t} \Delta s_{it} * (\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})}^{Between} \right) + \left(\overbrace{\sum_{i \in C_t} \Delta s_{it} * \Delta \boldsymbol{\varphi}_{it}}^{Covariance} \right) +$$

$$\left(\overbrace{\sum_{i \in EN_t} s_{it} * (\boldsymbol{\varphi}_{it} - \boldsymbol{\varphi}_{t-1}) - \sum_{i \in EX_t} s_{it} * (\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})}^{Net\ Entry} \right), \tag{4}$$

where the 'Within' and 'Covariance' terms are identical to those calculated using the Baily *et al*. (1992) method. The rest of the other ALP growth components calculated using Foster *et al*. (2001) are described as

Between-plant effects: $\sum_{i \in C_t} \Delta s_{it} * (\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})$ is the sum of the changing labour shares weighted by the deviation of initial plant-level productivity from initial industry productivity index. An increase in a continuing plant's labour share makes a positive contribution to the 'Between' component only if its initial productivity exceeds the average initial industry productivity.

Net-entry effects: The 'Entry' term, $\{s_{it}(\boldsymbol{\varphi}_{it} - \boldsymbol{\varphi}_{t-1})\}$, reflects the deviation of current firm-level productivity from average initial industry productivity index weighted by current labour shares. First, a new firm contributes positively to growth if its productivity level exceeds the average initial industry productivity index; i.e., $\boldsymbol{\varphi}_{it} > \boldsymbol{\varphi}_{t-1}$. Second, the 'Exit' component is calculated similarly to the 'Between' term, except that it is weighted by the un-differenced labour shares. Thus, a shutting down plant contributes positively to ALP growth only if it has lower productivity than the average initial industry productivity index; i.e., $(\boldsymbol{\varphi}_{it-1} < \boldsymbol{\varphi}_{t-1})$.

3.4.4   The Relationship Between the Baily *et al*. (1992) and Forster *et al*. (2001) Methods

The last two sections have outlined and discussed methods of decomposing the ALP index based on Baily *et al*. (1992) and Forster *et al*. (2001) but do not address their differences in calculating and

---

[30] King and Nielson (2016) argue in the context of propensity score matching that standardization of variables makes the analysis invariant to the substance.

interpreting the 'Between' and 'Net-Entry' components. In the examination of these methods, the scrutiny of the first and third terms in Eqs.3-4 is not undertaken because these terms are not model dependent. That is, these components are identical regardless of the model used to compute productivity gains. Therefore this sub-section considers the relationship between these methods and offers an explanation of the meaning of results thus generated.

The discussion of how the Baily *et al*. (1992) and Forster *et al*. (2001) approaches are related is best expressed mathematically as

$$FHK_{Bet} - BHC_{Bet} = BHC_{Net-Entry} - FHK_{Net-Entry}$$

The left-hand side of the expression relates the FHK_{Bet} to the BHC_{Bet} quantity for continuing plants. The latter is just a change in labour shares, weighted by an initial firm-level productivity that is always positive. This is a between-firm index measuring the productivity-weighted share shifting effects of a change in labour. The between-effects of the Baily *et al*. (1992) method can in principle either be positive due to labour growth, zero due to firm size stagnation or negative due to a producer scaling down operations. However, as noted by Haltiwanger (1997), the absence of a relationship between the initial firm-level productivity and initial industry average productivity does not guarantee a zero outcome in the between-firm effects index, even if all plants have the same productivity levels across the $t-1$ and t periods. In the case of the first term on the left-hand side, the weighting is based on deviations between the initial firm-level and average initial industry-level productivities. Unlike the Baily *et al*. (1992), the Forster *et al*. (2001) method therefore allows the weighting index to be positive if the initial firm-level productivity is lower than the industry average, zero if the initial firm-level and initial industry average are equal or negative if the initial firm-level productivity is lower than the industry average product.

Since the labour change across methods can take any sign while the productivity weight in BHC_{Bet} is always positive, FHK_{Bet} and BHC_{Bet} can have opposite signs and differing orders of magnitude for at least two reasons. First, assume a firm is hit by a negative exogenous shock and is forced to scale down operations by reducing its industry share of employment; i.e., $\frac{L_{it-1}}{L_t} \rightarrow \frac{L_{it}}{L_t}$, holding L_t constant in both periods. Since $L_{it-1} > L_{it}$, then the change in the firm's labour share at time t is $\Delta s_{it} = \frac{L_{it}}{L_t} - \frac{L_{it-1}}{L_t} < 0$. Given that the ratio of real value-added to labour, $\boldsymbol{\varphi}_{it-1}$, is always positive, then firm i's $BHC_{Bet} = \Delta s_{it} * \boldsymbol{\varphi}_{it-1}$ is negative, suggesting a movement of labour from the downsizing firm to other producers. If the same firm operated at lower efficiency levels than the initial average industry productivity index; that is, $\boldsymbol{\varphi}_{it-1} < \boldsymbol{\varphi}_{t-1}$, then the firm's $FHK_{Bet} = \Delta s_{it}(\boldsymbol{\varphi}_{it-1} - \boldsymbol{\varphi}_{t-1})$ is positive. Only if $\boldsymbol{\varphi}_{it-1} > \boldsymbol{\varphi}_{t-1}$ does FHK_{Bet} become negative for this type of firm. Both measures of

'Between' effects jointly suggest that labour resources in inefficient downsizing firms reallocate to initially more productive firms relative to the initial industry average productivity.

Second, firm i may experience a large positive demand shock and raise its employment at time t by drawing workers (i.e., $\Delta s_{it} > 0$) from firm i' to increase its production. Although the Baily *et al.* (1992) 'Between' effects will be positive, the Forster *et al.* (2001) 'Between' effects will either be positive, zero or negative, depending on whether $\boldsymbol{\varphi}_{it-1} > \boldsymbol{\varphi}_{t-1}$, $\boldsymbol{\varphi}_{it-1} = \boldsymbol{\varphi}_{t-1}$ or $\boldsymbol{\varphi}_{it-1} < \boldsymbol{\varphi}_{t-1}$ which indicates the direction of resource flows. That is, if $\boldsymbol{\varphi}_{it-1} > \boldsymbol{\varphi}_{t-1}$, for example, labour is moving an initially high efficient firm to an initially inefficient industry average of firms. Otherwise, if $\boldsymbol{\varphi}_{it-1} < \boldsymbol{\varphi}_{t-1}$, labour resources reallocate to initially more productive firms.

On the right-hand side, the expression relates net-entry effects computed from $BHC_{Net-Entry}$ and from $FHK_{Net-Entry}$ indices. In the Bailey *et al.* (1992) approach, the net effect of entrants and exiting producers reflects any differences in the levels of productivity between firm birth and death, and any differences in labour shares. In particular, and holding labour shares of the entrant and exiting plants constant, the net-entry productivity index is negative if the existing firm is more productive than the new born. Again, the index can also be negative if the quitting firm has a larger share of labour in the industry than does the entrant, holding firm-level productivity constant. This productivity measure is positive if the existing firm is less productive than the new born, holding labour shares of the entrant and exiting plants constant. It can also be positive if the quitting firm has a lower share of labour in the industry than does the entrant, holding firm-level productivity constant. In the case of the Forster *et al.* (2001), net-entry effects of productivity are driven by weighted deviations of the firm-level productivity from the initial industry average productivity instead of just the firm-level ratio of real value-added to labour. Thus, net-entry is positive if the productivity contribution from entry is greater than the productivity contribution from exit. This can happen only if the entrant is more productive than the initial industry average productivity *and* the exiting plant is less productive than the initial industry average productivity. Otherwise, net-entry is either negative or zero.

3.4.5   A Detailed ALP Decomposition for the Swazi Manufacturing Sector

The previous sections have outlined and discussed the two traditional methods of aggregate labour productivity decomposition, highlighting the impact of specific firm-level patterns of resource shares and productivity either in isolation or relative to the industry average. That enquiry does not clarify with precision how the identified micro-factors interact to dominate in a broadly defined industry. This section is concerned with a detailed analysis of the Swazi manufacturing sector to gain insight into the annual patterns of productivity variation represented by cross-plant movement of resources, technical change as well as net-entry dynamics. It achieves this by using an unbalanced dataset of

heterogeneous producers across 13 two-digit ISIC industries in the period 1994−2003. The estimation of ALP and its component parts is based on the Baily *et al.* (1992)/Foster *et al.* (2001) decomposition in Eq. 3 and 4 and reported in Table 3.3.

Table 3.3: ALP growth rate in Swazi manufacturing 1994–2003: Baily *et al.* (1992)/Foster *et al.* (2001) Decomposition Using Eq. 3 and Eq. 4 for Columns 3−7.

year	Value-Added Growth	Labour productivity growth (0)	Baily *et al.* (1992) and Foster *et al.* (2001) ALP decomposition: (0) = (1) + (2) + (3) + (4)					
			Within (1)	Between (2)		Cross (3)	Net Entry (4)	
				BHC-RE	FHK-RE		BHC	FHK
1995	7.76	-26.03	-12.34	-27.06	0.67	3.60	9.51	-17.83
1996	23.10	-1.33	-7.93	-6.92	0.08	0.23	13.37	6.28
1997	-44.35	-2.79	2.93	32.90	-8.51	4.09	-42.71	-1.31
1998	265.55	119.30	1.32	-36.46	0.40	-0.90	155.33	118.47
1999	275.57	102.21	9.89	-35.03	5.11	-3.86	134.79	91.46
2000	-16.28	-17.27	-17.10	-1.39	-3.20	0.11	0.05	3.67
2001	37.42	-1.02	31.79	-28.89	-0.42	-8.90	4.97	-23.50
2002	-20.74	-39.18	-25.93	-21.55	-3.62	5.60	2.71	-15.22
2003	-36.71	-3.33	-26.55	73.21	41.25	-29.12	-20.88	11.09
Mean	**54.59**	**14.51**	**-4.88**	**-5.69**	**3.53**	**-3.24**	**28.57**	**19.23**
Median	**7.76**	**-2.79**	**-7.93**	**-21.55**	**0.08**	**0.11**	**4.97**	**3.67**
Std Dev	**125.32**	**56.26**	**18.67**	**36.72**	**14.63**	**10.66**	**68.48**	**50.43**

Notes: The "Labour productivity growth" column depicts the ALP growth with entry and exit, and the "Value-added growth' column represents the aggregate real value added growth rate. The plant-level real value added is summed and annualized across plants. As in Nishida *et al.* (2014), numbers are percentage growth rates. We define labour productivity as the amount of real value added relative to unit labour. $\Delta\boldsymbol{\varphi}_t$ is decomposed into four components: (1) within, (2) between, (3) cross, and (4) net-entry term, using Eq. 1 in text for Baily *et al.* (1992) and Eq. 2 in text for Foster *et al.* (2001). We use employment share for the share weights, and both "within" and "between" terms use the base-period share for the weights.

Source: Author's own calculations.

The second and third columns report annualized growth rates in real value-added and ALP, respectively. The annual average (median) growth rate in real value-added is 54.59 percent (7.76 percent) with the measured standard deviation of 125.32 percent. Although real value-added growth is largely positive, particularly in 1998 and 1999, the incidence of negative growth is non-negligible. ALP, on the other hand, had an annual average (median) growth rate of 14.51 percent (−2.79 percent). Again, the years 1998 and 1999 stand out as outliers.[31] In seven out of nine years, we observe negative ALP values in column three.

---

[31] We made an attempt to remove any potential outliers as in Nishida et al. (2014) by applying the Stata "**Winsor**" command to the plant-level labour productivity at $p(0.01)$, which specifies the proportion of observations to be modified in each tail. This creates too many missing values and therefore we abandoned the procedure. Another approach involves identifying outliers and removing them sequentially, beginning with the largest. When the very first outlier where $\boldsymbol{\varphi}_{it} = 1.6$ is removed, decompositions for both 1998 and 1999 disappear. Again, this procedure is abandoned. However, it is considered not fatal to use the data 'as is' given

In columns four through nine, we present the Baily *et al.* (1992) and Foster *et al.* (2001) decompositions. The annual average 'within-effect' in column four is −4.88 percent compared to the Baily *et al.* (1992) between-plants term of −5.69 percent and Foster *et al.* (2001) between-plants term of 3.53 percent. Clearly, real productivity dominates the Baily *et al.* (1992) share-shift component of aggregate productivity, yet it is subordinate to the Foster *et al.* (2001) between-plants term. However, if the potentially profound confounding effects of entry−exit dynamics in the measured "Between" term calculated using the Baily *et al.* (1992) approach is accounted for, then net-entry and the "Between" effects dominate the measured "Within" effects. Both Baily *et al.* (1992) and Foster *et al.* (2001) decompositions make significant net-entry contributions to ALP growth by contributing 28.57 percent and 19.23 percent, respectively. The entry of more productive firms than the average initial industry productivity and the exit of lower productivity firms than the average initial industry productivity are the main drivers of ALP.

Looking at firm-level production efficiency in isolation, we find evidence of progressive weakening of technical change in manufacturing potentially induced by increasing competition in the Customs Union, save for the 31.79 percent productivity increase in 2001 which was consistent with the start of AGOA. Judging from the size of the standard deviation, there was marked heterogeneity in plant-level technical efficiency around a declining average productivity trend.

In a closer examination of incumbents, entrants and exiting firms, we find evidence of significant heterogeneity as in Liu and Tybout (1996) represented by the standard deviations of 68.48 percent and 50.43 percent in the Baily *et al.* (1992 and the Foster *et al.* (2001) approaches, respectively. We also find that, on average, exiting plants are 28.97 percent and 19.23 percent lower than incumbents in terms of productivity contribution to ALP when using the respective methods. Hence, their disappearance improves sectoral productivity. However, the occasional exit of relatively more efficient firms has the consequence of inducing a negative turnover effect on aggregate labour productivity. In this context, Liu and Tybout (1996) note that while productivity of exiting firms may drop, surviving entrants may raise their productivity such that the snowballing effects of this cleansing process are probably substantial over a longer time horizon. According to Caballero and Hammour (1994), it is this continuous process of creation and destruction of business units resulting from product and process innovation that is essential for understanding growth.

A further isolation of incumbents shows that productivity heterogeneity remains important, regardless of the approach used. Using the Forster *et al*. (2001) approach, we find the portion of change in sectoral productivity that is due to the labour market share reallocation accounts for 3.53 percent, on

---

the large similarities between our results and the results found in the literature, and the fact that the Swazi manufacturing sector is highly concentrated and these are real and important firms.

average. As in Nishida *et al.* (2014), it is instructive to determine the impact of an expanding or shrinking economy on the Baily *et al.* (1992) share-shift component. The direction of change in the number of firms can work to reduce or increase this component of productivity, as shown in the next section.

3.4.6   Confounding Effects of Firm Turnover on the Baily *et al.* (1992) Reallocation

The Baily *et al.* (1992) reallocation component can be further decomposed into two more constituent parts: one related to reallocation and another related to the number of plants as in Nishida *et al.* (2014). Suppose there are N_t plants in manufacturing at time *t* and the plant-level average share of employment is $s_t = \frac{\sum_i s_{it}}{N_t} = \frac{1}{N_t}$. Then, the relative labour share in the ith plant is defined as $\tilde{s}_{it} = s_{it} - s_t$, and the change in the relative labour share from time *t-1* to *t* is $\Delta \tilde{s}_{it} = \tilde{s}_{it} - \tilde{s}_t$. Hence, the "Between" term for incumbent firms can be decomposed as follows:

$$BHC_{RE} = \left(\sum_{i \in C_t} \Delta s_{it} * \boldsymbol{\varphi}_{it-1} \right)$$

$$= \left(\sum_{i \in C_t} \{ (s_{it} - s_t) - (s_{it-1} - s_{t-1}) \} * \boldsymbol{\varphi}_{it-1} \right) + \left(\sum_{i \in C_t} (s_t - s_{t-1}) \sum_{i \in C_t} \boldsymbol{\varphi}_{it-1} \right)$$

$$= \left(\overbrace{\sum_{i \in C_t} \Delta \widetilde{s_{it}} * \boldsymbol{\varphi}_{it-1}}^{First\ Component} \right) + \left(\overbrace{\left(\frac{1}{N_t} - \frac{1}{N_{t-1}} \right) \sum_{i \in C_t} \boldsymbol{\varphi}_{it-1}}^{Second\ Component} \right) \tag{5}$$

where C_t refers to continuing plants at time *t*. The first component represents labour reallocation and the second component is related to patterns of creative destruction. An increase in the number of firms over time confounds the first component by $\left(\frac{1}{N_t} - \frac{1}{N_{t-1}} \right)$ in the negative direction, since $\boldsymbol{\varphi}_{it-1}$ can never be negative. The reverse effect obtains in case of a persistent fall in the number of firms. The second component also gets smaller and smaller as the number of firms gets smaller and smaller, which happens if firm exit rate is persistently higher than the entry rate. If there is no change in the number of firms in the adjacent periods, the second component falls away. That is, the entry-exit dynamics have a spurious influence on the Baily *et al.* (1992) labour reallocation effect. Table 3.4 presents a quantitative decomposition of BHC_{RE} for the Swazi manufacturing sector.

Table 3.4: The ALP Growth Rate for the Swazi Manufacturing Sector (1994–2003): Baily *et al*. (1992) Between Term Decomposition.

Year	BHC (0): Between	Baily *et al*. (1992) between term decomposition: (0) = (1) + (2)		Percentage Growth of firms
		(1) First component	(2) Second component	
1995	-27.06	-16.93	-10.13	11.11
1996	-6.92	4.71	-11.63	13.75
1997	32.90	30.89	2.00	-2.20
1998	-36.46	-13.21	-23.25	25.84
1999	-35.03	-24.83	-10.20	23.21
2000	-1.39	4.08	-5.47	7.97
2001	-28.89	-21.39	-7.50	10.07
2002	-21.55	-14.49	-7.06	8.54
2003	73.21	54.59	18.62	-15.17
Mean	**-5.69**	**0.38**	**-6.07**	**9.24**
Median	**-21.55**	**-13.21**	**-7.50**	**10.07**
Std Dev	**36.72**	**26.73**	**11.39**	**12.37**

Notes Percentage growth rates. The Baily *et al*. (1992) 'between' term is decomposed into two terms using Eq. 5 in the text.

 Source: Author's own calculations.

The second column is identical to the BHC_{RE} column in Table 3.3 in the previous section. The third and fourth columns are the respective first and second components of Eq. 5, and the last column is the percentage growth of firms per year. In seven out of nine years, the manufacturing sector experienced growth in the number of firms, and in these years the confounding effect of plant expansion was negative on the 'Between' term. The comparison of the first term to the overall average of the Baily *et al*. (1992) 'Between' term shows that on average it is 6.07 percent higher over the sample period due to the downward confounding effects of plant turnover on the labour reallocation component. These results mimic the findings by Nishida *et al*. (2014) for Chile and Slovenia, and they cast doubt on the validity of the share-shifting effects of the Baily *et al*. (1992) approach. This confirms the conclusion by Nishida *et al*. (2014) that the Baily *et al*. (1992) reallocation can be negatively correlated, positively correlated or simply uncorrelated with the actual reallocation of inputs. A crucial argument in that paper, also corroborated by our results, is that the Baily *et al.* (1992) indices can erroneously equate reallocation growth to productivity growth, yet output per labour ratio is neither a perfect proxy for marginal products nor plant-level productivity.

This dilemma opens a door to the application of one of the promising approaches to estimating the decomposition of APG based on parametric aggregation of plant-level productivity. In his study of the robustness of productivity estimates, Van Biesebroeck (2007) demonstrates with Monte Carlo techniques the circumstances in which each of the methodologies works well. Among the six approaches analysed, two parametric methods appear suited to investigating productivity growth; namely, the systems generalized method of moments' estimator (SYS-GMM) and the semiparametric Olley and Pakes (1996)/Levinsohn and Petrin (2003)-type models.

The next sections draw heavily on the theoretical foundations of Petrin and Levinsohn (2012) as applied in Nishida *et al.* (2014) for measuring APG using plant-level data. Our purpose is to estimate and contrast the APG sources with those found when using traditional methods. It begins by determining a suitable proxy for the unobserved firm-level productivity. The actual semiparametric model estimation follows immediately.

3.4.7   Country Comparison of Evidence on Drivers of ALP Growth

In this section, the empirical decomposition of $\Delta\boldsymbol{\varphi}_t$ into its component sources of growth is reviewed for other countries for comparative examination. Two meta-analyses by BHS and by PPS together analyse 25 countries across Europe, the Americas and East Asia. Isaksson (2010) also surveys sources of $\Delta\boldsymbol{\varphi}_t$ in 33 advanced and developing countries as well as economies in transition, which include many of the countries covered in the BHS/PPS meta-analyses. A number of these countries have undergone economic reforms to facilitate freer movement of inputs across firms in order to trigger productivity growth from resource reallocation. A consistent finding is that there has been significant ALP growth, measured as growth in $\boldsymbol{\varphi}_t = \frac{\sum_i VA_{it}}{\sum_i L_{it}}$, for these economies.

In order to examine the sources of ALP growth, the BHS/PPS meta-studies decompose this index into real productivity and reallocation terms using the Baily *et al.* (1992) and Foster *et al.* (2001) methods. The survey by Isaksson (2010) adds Haltiwanger (1997) in its arsenal of techniques of productivity decomposition.[32] A key finding is that most of the growth in aggregate labour productivity comes from longitudinal firm-level efficiency gains; that is, 'Within' dominate 'Between' effects. Specifically, nine of the 25 countries experienced *negative* growth from resource reallocation and only four had a weak 'Between' term. Furthermore, 23 of the 25 countries had a negative covariance term.

Table 3.5 presents empirical decompositions of $\Delta\boldsymbol{\varphi}_t$ for the manufacturing sector covering a sample of 13 countries from the survey by Isaksson (2010), *plus* Swaziland, based on either the Foster *et al.* (2001) or Haltiwanger (1997) methods. This allows us to compare the results from Swaziland with evidence from market economies, economies in transition and Sub-Saharan Africa (SSA). Following the example of Van Biesebroeck (2005) for the Sub-Saharan results, we estimate a value-added production function which enables comparison of our results with those of other Sub-Saharan economies. Unlike Van Biesebroeck (2005), however, we also calculate productivity contributions

---

[32] The difference between Baily *et al.* (1992) and Haltiwanger (1997) is that the latter introduces the covariance term.

coming from entry and exit of firms, which now enables comparison with results from advanced nations and economies in transition.[33]

Table 3.5: ALP Growth, $\Delta\varphi_t$, Decomposition for the Manufacturing Sector in Industrialized Countries, Economies in Transition and in Developing Countries (Percentage) using Eq. 4.

Method	Country	Period	Output/Share/ Productivity	Within	Between	Cross	Entry	Exit	Total
FHK (2001)	USA	1992 & 1997	GO/Labour/LP	109.00	-3.00	-24.00	-29.00	49.00	102.00
FHK (2001)	UK	2000-2001	GO/Labour/LP	48.00	19.00	-17.00	35.00	12.00	97.00
FHK (2001)	Germany	1993-2003	GO/Labour/LP	118.60	11.50	-30.10	–	–	100.00
FHK (2001)	Russia	1992-2004	GO/Labour/LP	-590.40	359.60	61.61	-223.70	292.93	-99.96
FHK (2001)	Slovenia	1997-2001	GO/Labour/LP	68.00	18.00	-2.00	15.00	13.00	112.00
FHK (2001)	Chile	1985-1999	GO/Labour/LP	95.00	25.00	-50.00	-35.00	65.00	100.00
FHK (2001)	Colombia	1987-1998	GO/Labour/LP	105.00	20.00	-45.00	-20.00	40.00	100.00
FHK (2001)	**Swaziland**	**1994-2003**	**VA/Labour/LP**	**-33.63**	**24.33**	**-22.33**	**116.20**	**16.33**	**100.90**
Halti (1997)	Cameron	1990-1995	VA/Labour/LP	144.94	-25.84	-13.48	–	–	105.62
Halti (1997)	Ghana	1990-1995	VA/Labour/LP	78.97	66.15	-43.59	–	–	101.53
Halti (1997)	Kenya	1990-1995	VA/Labour/LP	445.45	282.80	-629.09	–	–	99.16
Halti (1997)	Tanzania	1990-1995	VA/Labour/LP	122.00	13.00	-36.00	–	–	99.00
Halti (1997)	Zambia	1990-1995	VA/Labour/LP	357.14	28.57	-278.57	–	–	107.14
Halti (1997)	Zimbabwe	1990-1995	VA/Labour/LP	163.33	33.33	-96.67	–	–	99.99

Notes: Methods are described in the text. LP = Labour Productivity, GO = Gross Output, VA= Value Added, and Halti (1997) = Haltiwanger (1997). Information sources include Isaksson (2010), "Structural Change and Productivity Growth: A Review with Implications for Developing Countries", *United Nations Industrial Development Organization,* Tables 1-3; Van Biesebroeck (2005), "Firm Size Matters: Growth and Productivity Growth in African Manufacturing", *Economic Development and Cultural Change*, Vol. 53(3), pp. 543-83; and the author's calculation of ALP growth components for Swaziland.

The average industry productivity for non-SSA countries, excluding Russia, is 101.83 percent and for SSA excluding Swaziland is 102.07 percent. This compares with 100.90 percent for Swazi manufacturing. The 'Within' effects generate more growth than 'Between' effects across all countries except Swaziland. In 12 of 14 countries, results show dominance of real productivity over both resource reallocation among incumbents and turnover effects. Sub-Saharan 'Within' effects also dominate share-shift effects in the rest of the other economies surveyed in the table. This suggests that the Sub-Saharan manufacturing sectors generate incredibly more productivity growth from innovation and technological progress than do the more technologically advanced economies. The highest beneficiary from technological advancement is, for example, Kenya with 445 percent 'Within' effects followed by Zambia with 357 percent. On the other hand, looking at the 'Between' term alone shows that only the U.S. and Cameroon had negative growth. Contrary to normally functioning market economies, this suggests that the U.S. manufacturing sector reallocated resources from high- to low-productivity plants between 1992 and 1997; and Cameroon did the same in the period 1990 to 1995.

Finally, while all countries reporting on turnover have positive growth from firm exit, only Swaziland, Slovenia and the UK report positive entry contributions to growth. The 16.33 percent for Swaziland

---

[33] Van Biesebroeck (2005) uses data from the RPED surveys of the World Bank spanning a maximum of five years for each country.

means that the country experienced the exit of lower productivity firms than the average initial industry productivity index. At the same time, Swaziland also experienced firm entry with higher average productivity of 116.2 percent than the average initial industry productivity. It can be shown that an un-normalized entry−exit rationalization effect of firms has a pronounced impact of 19.23 percent on ALP growth in Swaziland.

Moreover, the stylized fact from BHS/PPS and Isaksson (2010) is that real productivity dominates both the share-shift effects and turnover terms in studies that use Baily *et al.* (1992) or its derivatives such as Foster *et al.* (2001) and Haltiwanger (1997).[34] Contrary to conventional wisdom, however, the Swazi results show superiority of resource reallocation among incumbents and firm entry-exit dynamics over real productivity. This suggests that the Swazi manufacturing sector is unique in delivering dominance of reallocation and rationalization effects over innovation and technological advancement during a period of trade reforms.

3.5   The Petrin-Levinsohn (2012) Approach to Aggregate Productivity Growth Decomposition

3.5.1   Production Function Specification

The estimation of production functions in economics has been a fundamental activity in applied economics since the 1800s, and the early econometric problems inhibiting efficient estimation of the coefficients of capital and labour are still a concern even today. Perhaps the most recurring econometric issue is the likelihood of the presence of output determinants that are unobserved to the analyst but observed by the producer. If that is the case, and if capital and labour are chosen as a function of these output determinants, then there exists an endogeneity problem. In such situations, the OLS procedure generates biased parameters for the observed production inputs; see Ackerberg *et al*. (2015).

The semiparametric method of estimating production functions initiated by Olley and Pakes (1996) addresses problems of endogeneity in inputs and the unobserved productivity shocks. Instead of using

---

[34] What also stands out as a stylized fact from this analysis is that the sources of growth for ALP differ by country, period in a country and methodology applied to the sector in question. For example, in their analysis of the manufacturing sector in 1995−2000 as opposed to 1997−2001 above, De Loecker and Konings (2006) use the Foster *et al.* (2001) decomposition of ALP and find 'within' firm productivity growth of 123.4 percent and reallocation growth of -11.7 percent compared to 68 percent and 18 percent above, respectively. Simply by discarding the first two years and the last year of study, significantly different results are produced; see note 5 in Nishida et al. (2014) for the case of Chile, Colombia and Slovenia.

lumpy investment as a proxy for productivity like Olley and Pakes (1996), the Levinsohn and Petrin (2003) approach uses the intermediate input to estimate the gross output production function [35]

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \nu_{it}, \tag{6}$$

where all variables are in natural logarithms. The variable y_{it} is real output, β_0 is the constant term, the coefficients (β_l, β_k) are y_{it} elasticities with respect to labour and capital inputs.[36] l_{it} is variable labour input for firm i at time t, k_{it} is fixed and/or quasi-fixed capital input. The last two components are the unobservable productivity, ω_{it}, which is known to the firm but unknown to the econometrician, and ν_{it} is a sequence of independent and identically distributed (i.i.d.) shocks. Demand for intermediate inputs, m_{it}, is a function of state variables k_{it} and ω_{it} and is assumed monotonically increasing in ω_{it}. Therefore, this function is invertible to express ω_{it} as a function of k_{it} and m_{it}. In turn, ω_{it} is governed by a first-order Markov process with an additional innovation that is uncorrelated with k_{it}, but not necessarily with l_{it}.

In the first stage, Levinsohn and Petrin (2003) transform (6) into a function of labour input and an unknown function $g(k_{it}, m_{it})$, where $g(.)$ is approximated with a third-degree polynomial in k_{it} and m_{it}, and β_l is estimated using O.L.S., see Eqs. $1.6 - 1.8$ in Appendix A3.1. Ackerberg *et al.* (2015) (hereafter referred to as ACF) demonstrate how β_l is unidentified because l_{it} is correlated with $g(.)$, and propose an alternative but still two-stage approach. The second stage in Levinsohn and Petrin (2003) involves nonparametric estimation of the value of $\hat{\phi}_{it} = \hat{y}_{it} - \hat{\beta}_l l_{it}$, and estimating the productivity series using $\widehat{\omega}_{it} = \hat{\phi}_{it} - \beta_k^* k_{it}$. A consistent nonparametric approximation to $E(\omega_{it}|\omega_{it-1})$ is then given by predicted values from a nonlinear regression shown by Eq. 1.21 in Appendix A3.1. Given $E(\widehat{\omega_{it}|\omega_{it-1}})$, $\hat{\beta}_l$ and β_k^*, the estimate of β_k, solves the minimization of the squared regression residuals

$$\min_{\beta_k^*} \sum_i \left(y_{it} - \hat{\beta}_l l_{it} - \widehat{\beta_k^*} k_{it} - E(\widehat{\omega_{it}|\omega_{it-1}}) \right)^2. \tag{7}$$

Instead of a two-step approach, Wooldridge (2009) proposes to simultaneously estimate (β_l, β_k) by making a Conditional Mean Independence (CMI) assumption about the error term in respect of current and past values of l_{it}, k_{it}, m_{it}. This allows him to express the third-degree polynomial in single-period lags of capital and intermediate inputs as in (8)

---

[35] Appendix A3.2 shows 100 percent of non-zero intermediate observations compared to an average of only 34 percent for investment. Therefore, choosing investment as a proxy in this case would truncate 66 percent of the observations in the panel dataset.

[36] A detailed exposition of the Wooldridge-Levinsohn-Petrin estimation of the production parameters is found in Appendix A3.1.

$$y_{it} = \varphi_0^* + \beta_l l_{it} + \beta_k k_{it} + g(k_{it-1}, m_{it-1}) + u_{it} \tag{8}$$

or

$$y_{it} \cong \varphi_0^* + \beta_l l_{it} + \beta_k k_{it} + \sum_p^3 \sum_q^{3-p} \delta_{pq} k_{it-1}^p m_{it-1}^q + u_{it}. \tag{9}$$

Following Petrin and Levinsohn (2012), Petrin *et al*. (2011) and Nishida *et al*. (2014), Eq. 9 can be estimated using a pooled IV, with $k_{it}, k_{it-1}, l_{it-1}, m_{it-1}$ and third-order polynomial approximation of $g(.)$ with k_{it-1}, m_{it-1} as instruments for $l_{it} = f(l_{it}^{WP}, l_{it}^{PE})$, where l_{it}^{WP} denotes Working Proprietors and l_{it}^{PE} denotes Paid Workers. CMI II in the Appendix renders this approach robust to the ACF critique and it does not require bootstrapping to obtain robust standard errors for (β_l, β_k).

3.5.2   Parametric Estimation of the Production Function

It is essential to show in a practical sense how to efficiently estimate the parameters (β_l, β_k) using firm-level datasets. Eq. 9 can be estimated either by gross output production functions as in Petrin *et al*. (2011) or a value-added production technology as in Nishida *et al*. (2014). The latter is adopted here. Table 3.6 presents the characteristics of the empirical model.

Table 3.6: Specification of the Empirical Model

Panel A: Variables for the Levinsohn and Petrin (2003) or the LP Models	
Dependent variable:	Double-deflated value-added (rva_{it})
Freely variable inputs:	l_{it}^{WP}, l_{it}^{PE}
Proxy: Intermediate Inputs	m_{it}
Capital:	k_{it}
value-added:	$valueadded_{it}$
Reps (#):	Number of bootstrap replications to be performed
Panel B: Variables for the Wooldridge (2009) and Levinsohn and Petrin (2003) or the WLP Models	
Dependent variable:	Double-deflated real value-added (rva_t)
Included Instruments:	$k_{it}, k_{it-1}, m_{it-1}, k_{it-1}m_{it-1}, k_{t-1}^2, m_{t-1}^2, k_{t-1}^2 m_{it-1}, k_{it-1}m_{t-1}^2, k_{t-1}^3, m_{t-1}^3$
Endogenous variables:	l_{it}^{WP}, l_{it}^{PE}
Excluded Instruments:	$l_{it-1}^{WP}, l_{it-1}^{PE}$

Notes:
- Consistent with order conditions for identification in Hayashi (2000), the number of predetermined variables excluded from the equation $(l_{it-1}^{WP}, l_{it-1}^{PE})$ = the number of endogenous variables (l_{it}^{WP}, l_{it}^{PE}) or the number of instruments = the number of regressors.
- The test for weak instruments (Z variables) is H_0: Z$\in \mathcal{W}_{bias.TSLS}$ against H_1: Z$\notin \mathcal{W}_{bias.TSLS}$. The test procedure is, Reject H_0 if the Cragg-Donald (1993) g_{min} statistic $\geq d_{bias.TSLS}$(b;K_2, n, δ), where $d_{bias.TSLS}$ denotes the Stock and Yogo (2005) critical value based on the Two Stage Least Squares (TSLS) bias, K_2 the number of instruments, and n is the number of included endogenous regressors.

Panel A is the LP Model which includes freely variable inputs (l_{it}^{WP}, l_{it}^{PE}) and excludes the proxy variable m_{it}. ACF have however shown that the LP Model suffers from parametric identification problems arising from firms' optimization of variable labour, yet labour is also a deterministic function of unobservable productivity and capital. In Panel B, Wooldridge (2009) therefore modifies

Levinsohn and Petrin (2003) to correct for this endogeneity problem. In this Model, endogenous variables (l_{it}^{WP}, l_{it}^{PE}) are instrumented with capital and the polynomial approximation of the unknown expression $g(.)$.[37]

Table 3.7: Estimates of Production Functions with Third Order Polynomial

Variable	WLP	LP	FE[a.]	FE-Int[b.]	O.L.S.	O.L.S.lab
l_t^{WP}	-0.162	-0.118	-0.069	-0.028	-0.070	
l_t^{PE}	0.892***	0.794***	0.796***	0.793***	0.811***	
l_t						0.863***
k_t	0.224***	0.181***	0.216***	0.222***	0.193***	0.158***
m_t			0.325***	0.321***	0.306***	0.356***
k_{t-1}	7.074*					
m_{t-1}	0.663					
$k_{t-1}m_{t-1}$	-0.682**					
k_{t-1}^2	-0.162					
m_{t-1}^2	0.293**					
$k_{t-1}^2 m_{t-1}$	0.010					
$k_{t-1}m_{t-1}^2$	0.014*					
k_{t-1}^3	0.001					
m_{t-1}^3	-0.011***					
cons	-25.021		8.413***	8.397***	8.810***	8.367***
N	757	1021	1021	1021	1021	1257
R^2	0.839		0.811	0.827	0.796	0.824
R^2_a	0.837		0.806	0.803	0.795	0.824

Diagnostic Tests for the WLP Model

Endog Vars[c]	*Shea Partial R^2*	*Partial R^2*	*F(2,744)*	*p-value*
l_t^{WP}	0.3080	0.3219	41.69	0.0000
l_t^{PE}	0.8921	0.9324	3663.48	0.0000

[d.]Anderson-Rubin (AR) Test F(2,744)=172.86		0.0000	
Anderson-Rubin (AR) Test $\chi^2 = 351.77$		0.0000	
Stock-Write s Statistic $\chi^2 = 57.64$		0.0000	
[e.]Cragg-Donald (N-L)*CDEV/L1		F-Statistic =	165.59

Legend: * p<0.05; ** p<0.01; *** p<0.001.

Notes:
a. Represents a fixed effects' model that controls for both time and industry effects.
b. Represents a fixed effects' model that interacts time with industry effects.
c. The Shea (1997) partial R^2 provides evidence for the presence of significant correlation between excluded variables $\left(l_{it-1}^{WP}, l_{it-1}^{PE}\right)$ and endogenous regressors $\left(l_{it}^{WP}, l_{it}^{PE}\right)$.
d. H_0: $B1 = 0$ and overidentifying restrictions are valid. The null is strongly rejected by AR F- and χ^2- tests as well as by Stock and Write (2000) χ^2-test, where B1=0 tests the joint significance of coefficients of endogenous variables. See Stock and Yogo (2005) for a detailed and fairly accessible discussion.
e. H_0: instruments are weak, even though parameters are identified. The null is strongly rejected at 95% confidence when the statistic g_{min} =165.59 is compared with the TSLS critical value of 7.03 produced by K_2=13, n=2 and the desired maximum level of bias of the IV estimator relative to OLS bias (b)=10% as in Stock and Yogo (2005, table 5.1).

Table 3.7 presents estimation results from the WLP Model, Levinsohn and Petrin (2003), Fixed Effects and O.L.S. methods with separate and combined labour components. Our preferred production

---

[37] A full derivation of the empirical LP Model and its transformation into WLP Model is presented in Appendix A3.5.

function specification is the WLP version of Eq.9 as outlined in Appendix A3.5.[38] While $\ln WP$ is negative and insignificant across specifications, $\ln PE$ and $\ln K$ are consistently positive and highly significant. The model is well-behaved and its primary input parameters are comparable to ACF input coefficients in Gandhi *et al.* (2016, table 1) for the cases of Colombia and Chile.

One important finding from our preferred the IV−GMM estimator presented as the WLP Model is that primary inputs in manufacturing deliver increasing returns to scale. This is potentially associated with import-competing industries whose output is likely to decline due to intensified foreign competition during the trade liberalization episode in the Customs Union.[39] The low value of the capital coefficient is typical in the literature and the cited cause for this is measurement error; see Levinsohn and Petrin (2003).[40] The IV−GMM labour coefficient shows an improvement of 10 percent compared to the other estimation methods. This can be attributed to efficiency gains in the GMM routine induced by the removal of selection and simultaneity biases. Industry effects on real value-added movements show a significant degree of heterogeneity whereby five of the 13 industries made insignificant contributions to output and the Apparel industry suffered a marked decline, particularly in 2001. Furthermore, there is no evidence of time effects in the first seven years and a significant decline began persistently in 2000 with marked negative effects in 2001 and 2003. The economic performance in the latter years coincides with heightened firm exit and the near-conclusion of progressive tariff-cuts in SACU.

3.5.3   General Set-Up, APG Decomposition and Estimation

There is already a growing view noted by Banerjee and Moll (2010), among others, that countries' underdevelopment may not only be an outcome of resource inadequacy, such as capital, skilled labour, entrepreneurship, or ideas, but also a result of the misuse or misallocation of available resources. Specifically, Banerjee and Duflo (2005); Jeong and Townsend (2007); Restuccia and Rogerson (2008, 2012); Hsieh and Klenow (2009); Bartelsman *et al.* (2004); and Alfaro, Charlton, and Kanczuk (2008) all argue that the scope of resource misallocation in developing economies is large enough to explain a significant gap in the aggregate productivity growth between advanced and

---

[38] The **ivreg2** Stata command with the GMM continuously updated estimator (**cue**) and **cluster** for each firm in order to generate efficient IV-GMM parametric estimates of the WLP functional specification was used

[39] The constant returns to scale in the other estimation methods is potentially induced by simultaneity and selection problems explained in detail in Wooldridge (2001).

[40] Galuščák and Lizal (2011) correct for measurement error in the capital series by running an O.L.S. on $k_{it} = \gamma_0 + \gamma_1 z_{1t} + \cdots + \gamma_N z_{Nt} + e_{Nt}$, where e_{Nt} is the i.i.d. measurement error, z_{it} are instruments and the predicted values of capital are $\hat{k}_t = k_t - e_t$. The estimation proceeds with linear approximation of the unknown function for consistency, and coefficient standard errors are derived non-parametrically through bootstrapping that reflects uncertainty in capital adjustment. Improvement in the capital input measurement to investigate industries' scale economies is left for future work.

poor countries. A similar argument is relevant if trade reforms identify industries that still need protection while trade liberalization in other industries deepens, as demonstrated by Edwards (2006) in the case of South Africa and, by extension, the rest of SACU.

Furthermore, there are also factors that move an economy away from the perfect competition setting such as input adjustment costs, hiring, firing and search costs, holdup and other contracting problems, taxes and subsidies, and markups. Examples of empirical evidence include Kambourov (2009) for firing costs in the case of Chile and Mexico, Aghion, Brown and Fedderke (2007) and Fedderke, Kularatne and Mariotti (2005) for markups in South Africa, and Petrin and Sivadasan (2013) for marginal product-marginal cost gaps in Chile. The finding of input misallocation suggests the presence of barriers to the movement of resources across heterogeneous production units. Similarly, firm-level heterogeneity in marginal products of inputs within industries in a country is also pronounced; see, for example Hsieh and Klenow (2009) for the case of India and China, Petrin and Sivadasan (2013) for Chile and Ho *et al.* (2014) for Ecuador. Ho *et al.* (2014), Petrin *et al*. (2011) and Nishida *et al*. (2014) rely on Petrin and Levinsohn (2012) to identify the relative role of technical efficiency improvement, the intensive and extensive margins. In response to the non-neoclassical frictions in developing economies, we also implement the Petrin and Levinsohn (2012) approach to estimate the extent of technical efficiency improvement and both margins of reallocation.

3.5.4   The General Set-Up

In this section we focus on the reallocation of primary inputs across, and the patterns of technical efficiency within, firms. The characterization of aggregate productivity growth in the absence of intermediate inputs takes the form

$$\left\{\sum_i \sum_k \left(P_i \frac{\partial Q_i}{\partial X_k} - W_{ik}\right) dX_{ik}\right\} + \left\{\sum_i P_i \frac{\partial Q_i}{\partial \omega_k} d\omega_i\right\} \tag{10}$$

where $\frac{\partial Q_i}{\partial X_k}$ is the partial derivative of output with respect to capital. We denote the price of output Q_i in establishment i as P_i, and W_i denotes the cost of labour. The change in the use of kth input quantity X_{ik} by firm i is dX_{ik}. The 'net output' remaining after deducting contributions by factor inputs to dQ_i is $d\omega_i$. Therefore, $\sum_i P_i \frac{\partial Q_i}{\partial \omega_k} d\omega_i$ represents gains from total technical efficiency changes, given $d\omega_i$. In Petrin and Levinsohn (2012, Lemma 1) and Petrin *et al*. (2011, Eq. 7), the

impact of a change in the kth input on a change in output is normalized as $\frac{\partial Q_i}{\partial \omega_k} = 1$ to transform the total technical efficiency changes into $\sum_i P_i d\omega_i$. [41]

Thus, Eq. 12 shows that the primary input reallocation is zero if dX_{ik}=0. This occurs if distortions or adjustment costs are so prohibitively high that inputs do not adjust and consequently do not reallocate across firms. Furthermore, under a perfectly operating factor input market, the VMP of each input is equal to its reward, $P_i \frac{\partial Q_i}{\partial X_k} = W_{ik}$. This means that factor inputs are continuously reallocated across plants in response to changes in economic conditions to maintain the VMP-price equality and no extra output gains can be realized from this reallocation; see Petrin and Levinsohn (2012).

3.5.5   APG Decomposition and Estimation

The decomposition of APG based on a double-deflation procedure for the value-added function, if it exists, is shown by Petrin and Levinsohn (2012) to be

$$\text{APG} = \sum_i D_i^v \, dln(VA_i) - \sum_i \sum_k s_{ik} \, dln \, X_{ik} \tag{11}$$

where the Domar-weight $\left(D_i^v = \frac{VA_i}{\sum_i VA_i}\right)$ is plant i's real value-added share. The two classes of labour are defined as $X_{PE} = L^{PE}$ and $X_{WP} = L^{WP}$, where L^{PE} refers to Paid Employees and L^{WP} refers to Working Proprietors (or Nonproduction Workers in Levinsohn and Petrin (2003)). The real value-added production function can then be written as

$$ln(VA_i) = \sum_k \epsilon_{ik}^v X_{ik} + ln\omega_i^v. \tag{12}$$

Eq. 12 can be translated into APG as

---

[41] The definition of APG allows for the classification of firms into entrants and exits, and exporters and nonexporters. It is also flexible to account for the impact of growth of both firm-level fixed and sunk costs (F_i^v) and input (Capital, Labour, Energy, Material) reallocation contributions, see Bruno (1978, Section 3), Petrin and Levinsohn (2012:706) and Petrin *et al.* (2011, Eq. 10). This means Eq. 9 can be fully decomposed into the expression

$$\left\{\sum_i \sum_k \left(P_i \frac{\partial Q_i}{\partial X_k} - W_{ik}\right) d X_{ik}\right\} + \left\{\sum_i \sum_k \left(P_i \frac{\partial Q_i}{\partial M_k} - P_{ik}\right) d M_{ik}\right\} + \{\sum_i P_i d\omega_i\} - \{\sum_i P_i dF_i\}$$

where $\frac{\partial Q_i}{\partial \omega_k}$ in the third term is normalized to one and the expression is translated into an augmented version of APG in (12) as

$$\left(\overbrace{\sum_i D_i^v \sum_k (\varepsilon_{ik}^v - s_{ik}^v)\Delta ln X_{ik}}^{Primary\ Input\ Reallocation}\right) + \left(\overbrace{\sum_i D_i^v \sum_j (\varepsilon_{ij}^v - s_{ij}^v)\Delta ln M_{ij}}^{Material\ Input\ Reallocation}\right) + \left(\overbrace{\sum_i D_i^v \Delta ln\omega_i^v}^{Technical\ Efficiency}\right) -$$

$$\left(\overbrace{\sum_i D_i^v \Delta ln F_i^v}^{Fixed\ and\ Sunk\ Costs}\right).$$

$$APG = \left(\overbrace{\sum_i D_i^v \sum_k (\varepsilon_{ik}^v - s_{ik}^v) \Delta ln X_{ik}}^{Primary\ Input\ Reallocation} \right) + \left(\overbrace{\sum_i D_i^v \Delta ln \omega_i^v}^{Technical\ Efficiency} \right) \tag{13}$$

where the first-difference operator is $\Delta x_{it} = x_{it} - x_{it-1}$ and $s_{ik}^v = \frac{W_{ik} X_{ik}}{VA_i}$ is the kth input revenue ratio to the plant's real value added. The real value-added elasticity with respect to the kth input is $\varepsilon_{ik}^v = \frac{\varepsilon_{ik}}{1 - s_{ik}}$. The gaps in Eq. 13 are measured by the difference between the plant-level value-added elasticities (ε_{ik}^v) and its input revenue share (s_{ik}^v) to value added. The aggregate input reallocation is therefore given by $\sum_i D_i^v \sum_k (\varepsilon_{ik}^v - s_{ik}^v) \Delta ln X_{ik}$ and aggregate technical efficiency is $\sum_i D_i^v \Delta ln \omega_i^v$. The APG approach has been applied to the US manufacturing data by Petrin *et al*. (2011), to Chile, Colombia and Slovenia by Nishida *et al*. (2014), to Chile by Petrin and Sivadasan (2013) and to Ecuador by Ho *et al.* (2014).

Using index number theory, it is possible to estimate Eq. 10 directly from the discrete data using the Törnqvist−Divisia methods. As in Nishida *et al*. (2014), the prices in the Domar-weights are annually chain-weighted and updated. The Törnqvist−Divisia method can be used in Eq. 12 for each of the two APG components; namely, the reallocation of primary inputs and technical efficiency – the respective analogues to the 'Between' and 'Within' terms from ALP in the traditional approach. The *estimated* aggregate productivity growth can then be expressed as

$$\overline{APG}_{it} = \sum_i \overline{D}_{it}^v \Delta ln(VA_{it}) - \sum_i \overline{D}_{it}^v \sum_k \bar{s}_{it}^v \Delta ln X_{ikt}, \tag{14}$$

which translates to

$$\overline{APG}_{it} = \{\sum_i \overline{D}_{it}^v \sum_k (\varepsilon_{ik}^v - \bar{s}_{it}^v) \Delta ln X_{ikt}\} + \{\sum_i \overline{D}_{it}^v \Delta ln \omega_{it}^v\}. \tag{15}$$

The $\overline{D}_{it}^v$ denotes plant *i*'s average value-added share weight from year *t−1* to *t*, Δ the first difference operator as before, and $\bar{s}_{it}^v$ is the two-period average of plant *i*'s expenditure for the kth primary input as a share of firm-level value added. In summary, the definitions of the APG components are

Technical Efficiency: $\sum_i D_i^v \Delta ln \omega_i^v$ is the value-added production function sum of the Domar-weighted changes in the Solow residuals, the APG analogue of the ALP "Within" term in Baily *et al*. (1992)/Foster *et al*. (2001). Technical efficiency increases when a plant continuously innovates and adapts to technological advances through learning-by-doing/watching and other means.

Reallocation: $\sum_i \sum_k \left(P_i \frac{\partial Q_i}{\partial X_k} - W_{ik} \right) d X_{ik} \xrightarrow{yields} \sum_i D_i^v \sum_k (\varepsilon_{ik}^v - s_{ik}^v) \Delta ln X_{ik}$. According to Petrin and Levinsohn (2012), Petrin *et al*. (2011), Petrin and Sivadasan (2013, p. 288) and Nishida *et al*. (2013, Eqs. 6 and 8), plants produce at the output level where $P_i \frac{\partial Q_i}{\partial X_k} > W_{ik}$, under imperfect factor market

conditions. Therefore, there are three potential instances for input reallocation growth. *First*, if dX_{ik} is the change in the kth factor input that was previously idle, but now reallocates to plant *i*, then the value of aggregate output changes by $P_i \frac{\partial Q_i}{\partial X_k} - W_k$. *Second*, when a small amount of primary inputs reallocates from *j* to *i* so that $dX_i = -dX_j$, then aggregate output changes by $P_i \frac{\partial Q_i}{\partial X_k} - P_j \frac{\partial Q_i}{\partial X_k}$. *Third*, in the event factor inputs reallocate across firms but the total amount of these inputs is held constant, the change in aggregate output induced by reallocation is given by $P_i \frac{\partial Q_i}{\partial X_k} dX_{ik}$.

Entry and Exit: Entry in this set-up includes the development of a new product, the replication of an existing product by a new firm or a reintroduction of a good back into the market after exiting previously (see Petrin and Levinsohn, 2012: Appendix).

In order to separately estimate firm-level technical efficiency in Eq. 12 for each ISIC2-digit industry code in Swazi manufacturing, Eq. 6 can be re-written as

$$\widehat{ln\omega_{it}^v} = \{ln(VA_{it})\} - \left\{\widehat{\beta^v} + \widehat{\epsilon_{jPE}^v}lnL_{it}^{PE} + \widehat{\epsilon_{jWP}^v}lnL_{it}^{WP} + \widehat{\epsilon_{jWP}^v}lnK_{it}\right\} \qquad (16)$$

and estimated using the proxy method of Wooldridge (2009) that modifies Levinsohn and Petrin (2003) to address the simultaneity problem in the determination of inputs and productivity. In Eq. 14, we use three factor inputs as regressors: non-production (Working Proprietors) L_{it}^{WP}, production (Paid Employees) L_{it}^{PE} and capital K_{it}. Unlike Nishida *et al*. (2014), we do not report only aggregate labour reallocation in our results, we also report reallocation of L_{it}^{WP} and L_{it}^{PE} separately.

Table 3.8 quantitatively decomposes APG into technical efficiency, primary input reallocation and net entry estimated using Eq. 14. The relationship between APG and its component sources of growth is that APG(0) *equals* 'Technical Efficiency (1)' *plus* 'Total Reallocation (2)' *plus* 'Net-Entry (3)'. In turn, 'Labour Reallocation (2)' decomposes to 'Working Proprietors Reallocation' *plus* 'Paid Employees Reallocation' while 'Total Reallocation' refers to all primary input reallocation across plants. In considering the results sequentially, the second and third columns show changes in real value added and aggregate productivity, respectively. It is striking to observe such a high correlation between aggregate productivity growth and the growth of value added. This reflects the fact that most of the fluctuations in aggregate productivity are predominantly linked to fluctuations in value added. Similar results are found in the case of Chile, Colombia or Slovenia in Nishida *et al*. (2014) or for the case of Japan in Kwon *et al*. (2009). For example, the Swazi manufacturing sector reports an estimated average real value added of 54.59 percent and average APG of 54.54 percent, or the median real value added of 7.76 percent and the median APG of 7.71 percent per year, respectively.

Table 3.8: Aggregate multifactor productivity growth rate, Swaziland manufacturing 1994–2003: APG decomposition, manufacturing value-added index double-deflator.
Estimates of $\sum_i \overline{D}_{it}^v \Delta ln(VA_{it})$ and $\overline{APG} = \sum_i \sum_k \overline{D}_{it}^v (\varepsilon_{ik}^v - \overline{s}_{it}^v)\Delta lnX_{ik} + \sum_i \overline{D}_{it}^v \Delta ln\omega_i^v.$

Year	Value-Added Growth	APG (0)	APG Decomposition: (0)= (1) + (2) + (3)					Net Entry (3)
			Technical Efficiency (1)	Reallocation				
				Total Reallocation (2)	Labour Reallocation	Working Proprietors' Reallocation	Paid Employees' Reallocation	
1995	7.76	7.71	-4.43	-4.31	8.76	5.61	3.15	16.45
1996	23.10	23.03	2.27	-6.98	1.21	-0.30	1.51	27.75
1997	-44.35	-44.25	-2.69	18.13	9.84	-0.08	9.92	-59.69
1998	265.55	265.30	2.31	-2.38	0.10	0.02	0.07	265.37
1999	275.57	275.42	0.64	9.81	3.01	-0.01	3.03	264.97
2000	-16.28	-16.27	-15.30	-5.16	0.61	-0.13	0.74	4.18
2001	37.42	37.39	9.03	20.10	-0.08	-0.11	0.03	8.25
2002	-20.74	-20.75	-3.56	-29.01	-1.36	-1.66	0.30	11.82
2003	-36.71	-36.67	-20.74	1.12	7.14	-2.61	9.75	-17.05
Mean	**54.59**	**54.54**	**-3.61**	**0.15**	**3.25**	**0.08**	**3.17**	**58.01**
Median	**7.76**	**7.71**	**-2.69**	**-2.38**	**1.21**	**-0.11**	**1.51**	**11.82**
Std Dev	**125.32**	**125.23**	**9.21**	**14.88**	**4.22**	**2.27**	**3.96**	**120.14**

Notes: As in Nishida *et al*. (2014), numbers are percentage growth rates. The plant-level multifactor productivity is calculated by using production function parameters that vary across 2-digit ISIC. We obtain the estimates by using Wooldridge (2009). APG represents the aggregate productivity growth with entry and exit, which is defined as aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding input constant. We use value-added share (Domar) for weights. APG is decomposed into four components: (1) technical efficiency, (2) reallocation, and (3) net-entry term, using Eq. 17 in text.

These trends are characterized by high firm-level heterogeneity in the change of value added and APG. For example, the measure of dispersion for APG is over twice its average size. One channel explaining this is found in Syverson (2004), which states that trade liberalization creates a competitive market environment and snowballing of product variety. This enables consumers to switch between products and/or producers such that high-cost producers' profitability is diminished. Thus, a high substitutability industry is likely to have less productivity dispersion and a high aggregate productivity level.

The contribution of technical efficiency to APG is on average (median) −3.61 percent (−2.69 percent) per year, compared to an average of 0.95 percent for Chile, 0.25 percent for Colombia and 2.17 percent for Slovenia (see Nishida *et al*. (2014)). This component of APG is positive in only four out of nine years. However, the most interesting case is the combined input reallocation in the fourth column reflecting simultaneous cross-plant movements in capital and components of labour inputs. The average total reallocation is 0.15 percent per year and consists of input reallocation from low to high productivity plants, from idle state to productive uses and reallocation that is not accompanied by changes in input amounts. Clearly, the average reallocation compares with 1.60 percent for Chile, 3.63 percent for Colombia and 3.42 percent for Slovenia as reported in Nishida *et al*. (2014).

However, our ultimate focus is the behaviour of the paid labour resource in response to shifts in economic factors that cause movements in the manufacturing sector. We first isolate labour reallocation from the contribution of all inputs put together. This produces 3.25 percent as the average annual rate of labour reallocation, and we report only two instances of negative reallocation out of the nine years studied. A further decomposition of labour reallocation into that which is accounted for by the reshuffling of working proprietors and paid employees produces sharper results. Paid employment shows positive growth in every year and accounts for an average of about 98 percent [3.17 percent÷3.25 percent] of all labour reallocated per year. Again, paid labour reallocated from low to high VMP plants, new paid labour entered the labour market and some paid labour reallocated without increasing the number of workers. This is consistent with the wave of downsizing in the manufacturing sector during the period of trade liberalization. Our results are robust to the use of 'single-deflation' by the manufacturing value-added deflator in Appendix A3.3 and 'double-deflation' by the consumer price index in Appendix A3.4. Another robustness check applied, but not reported here, involved 'single-deflation' by the consumer price index which also sustained the basic results.

Thus, the analysis reveals that the contribution by the labour reallocation growth to APG decisively *dominates* technical efficiency in the manufacturing sector in Swaziland. Firms were not investing more in improving production efficiency through innovation and adoption of new technologies than they were moving labour to higher activity producers. This conclusion remains robust regardless of the deflation procedure used in the estimation of the value-added production function. However, based on our robustness checks, the combined input reallocation *versus* technical efficiency is inconclusive because the outcome depends on whether we use the mean or the median as a standard for comparison.

On the other hand, the extensive margin accounts for most of the change in APG. The annual average of net entry contribution to APG is 58.01 percent and is driven by the dramatic increase of APG in 1998 and 1999. This pattern of high contribution by net entry is consistent with extensive margin effects of trade liberalization which increases opportunities for mergers and acquisitions as well as business restructuring and retrenchments.

3.6   Discussion of Results

In the previous sections, different decomposition approaches for aggregate productivity growth are described, estimated and results compared. It is evident that the joint use of the Bailey *et al*. (1992) *and* Forster *et al*. (2001) methods to measure contributions made by individual determinants of the aggregate labour productivity growth produces significant insights. More specifically, while these methods identically define the longitudinal effects of productivity changes and the covariance effects, their conceptualization of resource-shift effects and the entry-exit dynamics differs only in terms of

whether or not firm-level productivity deviations from the initial industry average productivity is considered. That is, the Bailey *et al*. (1992) technique does not consider these deviations while Forster *et al*. (2001) does.

The present study of industrial aggregate productivity growth in Swaziland coincides with a period of progressive trade liberalization and deregulation in the customs union. Trade reforms typically create competitive markets by inducing domestic price reduction, forcing inefficient producers out of business thereby reallocating resources and market shares to more productive plants, see Pavcnik (2002). However, the standard absence of well-functioning markets due to other forms of protection in developing economies may account for the observed poor industrial performance in Swaziland. In table 3.3; for instance, the average year-on-year within-firm effects is negative. In five out of nine instances, within-effects report large negative productivity growth, suggesting that the manufacturing sector in was dominated by continuing low productivity firms. This productivity growth component is only positive in 1997-1999 and in 2001, suggesting the manufacturing sector in Swaziland experienced some productivity growth in these years. That is, the annual orders of magnitude in these specific years indicate that plant-level improvements in production efficiency only marginally dominated industrial activity. In an efficient market environment, the weak performance of the sector in technological advancements would feature prominently in heightened exit rates of poor performers and entry of efficient firms.

The labour share-shift effect computed from the traditional methods produces interesting results. On a year-to-year average basis, the Bailey *et al*. (1992) between-effect is -5.69 percent and the Forster *et al*. (2001) between-effect is 3.53 percent. Such patterns of negative Bailey *et al*. (1992) between-effects and positive Forster *et al*. (2001) between-effects occur in four out of nine instances. Interpreting these results collectively, it means most industrial firms downsized their operations and this affected mostly plants with initial productivity level that exceeded the initial industry average productivity. The observed apparent inefficient reshuffling of resources away from productive to less productive producers can be explained in terms of the newly reforming industrial sector in the customs union. These are likely South African owned subsidiaries that moved to Swaziland during the period of economic sanctions prior to the mid-1990s to access cheaper intermediate and primary inputs as well as foreign markets. The new trade policy regime was incentive enough for these plants to relocate back into the larger South African market to enjoy scale economies in an increasingly competitive market environment.

However, section 3.4.6 demonstrates that the traditional methods suffer from confounding effects of firm turnover. Purging these effects from the producer-level labour share merely reduced the magnitude of the share-shift effect in absolute terms without altering its sign and only converted this

effect from negative to positive in 1996. Therefore, our results generally remain robust to the confounding effects of changes in the number of firms over time.

The entry-exit dynamics that characterize the manufacturing sector in Swaziland tell an interesting story about the behavioural patterns of establishments when using the Bailey *et al*. (1992)/Forster *et al*. (2001) techniques during the 10-year period. Although both methods yield large positive net-entry effects of productivity growth on a year-to-year average basis, table 3.3 reports four out of nine instances of positive Bailey *et al*. (1992) net-entry effects associated with positive Forster *et al*. (2001) net-entry effects. Again, a joint interpretation of this result from the two methods is that new firms were generally more productive relative to *both* their exiting counterparts and initial industry average productivity. In turn, exiting plant productivity levels were predominantly lower than the initial industry average productivity. This pattern is more pronounced in 1998-1999, a period of significant shake up in one industry where a large investment asset was sold to another and this was recorded as firm entry. The results also show three out of nine instances of positive Bailey *et al*. (1992) net-entry effects associated with negative Forster *et al*. (2001) net-entry effects. This is evidence of more productive entrants than quitters, and more productive quitters than the initial industry average productivity.

The Bailey *et al*. (1992) approach and its associated derivatives has been fiercely criticised by Levinsohn and Petrin (1999), Petrin and Levinsohn (2012) and Petrin *et al*. (2014) for decomposing aggregate labour productivity growth using firm-level output per labour, $\boldsymbol{\varphi}_{it} = \frac{VA_{it}}{L_{it}}$, as a proxy for the marginal product of labour. This literature also questions the use of changes in output/labour, $\Delta\boldsymbol{\varphi}_{it}$, as a proxy for plant-level changes in productivity. Petrin *et al*. (2014) demonstrate *a priori* and in a firm-level panel data application to the cases of Chile, Colombia and Slovenia how plant specific technical efficiency, input reallocation and turnover effects influence changes in APG. Following this alternative line of enquiry into the behaviour of industrial determinants of APG in Swaziland, two technical activities are carried out. First, an analytical framework for estimating a robust production function for the thirteen two-digit ISIC industries is developed and implemented to understand the behaviour of capital and labour inputs in relation to real value-added. This exercise turned out crucial in the estimation of the Solow-residual for use in the subsequent analysis. Second, a conceptual framework based on Petrin and Levinsohn (2012) for estimating the impact of plant-level technical efficiency and resource reallocation across firms is outlined in full and applied to the manufacturing sector in Swaziland.

Table 3.8 presents results based on the Petrin and Levinsohn (2012)/Nishida *et al*. (2014) procedure for measuring technical efficiency, input reallocation and plant turnover effects on aggregate productivity growth. These results broadly mimic those generated from using the Bailey *et al*.

75

(1992)/Forster *et al*. (2001) methods. In Swaziland, the manufacturing sector is highly concentrated even within broadly defined industries. Hence, a major movement of resources between a few firms translates into significant output changes as observed in 1998-1999 of the second column. Since aggregate productivity growth is defined here as the change in aggregate final demand less the change in the aggregate expenditure in primary inputs, the measured aggregate productivity growth matches the industrial value-added growth very closely over time.

Technical efficiency is on average negative and annually traces the ALP within-firm effects produced by conventional methods closely, although the APG orders of magnitude are much lower in absolute terms. This confirms the earlier view that the degree of firm-level and industrial innovation and entrepreneurial transformation remains negligible at best in the period under study. The direct effect of the generally negative real productivity in Swaziland reverses any positive impact arising from other sources of AGP despite the unboundedness of learning and ingenuity opportunities available to firms as discussed in Levinhson and Petrin (1999).   Such preponderance of poor producer performance in a trade liberalization period associated with intensified import competition is hard to explain without thinking about a possible existence of protective industrial regulations, high costs of adjustment of primary inputs or managerial incapacity. Capital irreversibility and protective policies are a crucial barrier to firm exit. Evidence by Bloom *et al*. (2013) shows that the adoption of appropriate managerial practices in large Indian textile firms raised productivity by 17 percent in the first year.

The most important input of production to national policymakers, Bretton Woods institutions and development organizations in the context of Swaziland is paid labour employment. During the period of trade reforms, there was an average paid labour reallocation productivity growth of 3.17 percent every year. Looking at paid employee productivity that is in excess of one percent, this is observed only in five out of the 10 years. Three of these years experienced paid labour productivity that is *at most* 0.07 percent. Nonetheless, positive industrial paid labour reallocation characterized every single year. There are at least four explanations based on $VMP_{ik} = P_i \frac{\partial Q_i}{\partial X_k}$, value-added elasticities and input shares that shed some light into these patterns of growth. First, the reallocation of paid labour input from plant j to plant i leads to $dL_{iPE} = 1$ and $dL_{jPE} = -1$. This increases the amount of real value-added by

$$P_i \frac{\partial Q_i}{\partial L_{iPE}} - P_j \frac{\partial Q_j}{\partial L_{jPE}},$$

assuming common wages across firms and holding total labour input constant. Hence, when paid labour moves from low to high VMP_{iPE}, aggregate final demand increases without any increase in technical efficiency or aggregate input use, see Petrin and Levinsohn (2012).

Second, market distortions arising from markups and taxes, and the impact of adjustment costs of paid labour, find full expression in the resource reallocation component of APG. The markup is by definition the wedge between the price and marginal cost of the product in question, and APG increases when paid labour moves from low to high markup firms. On the other hand, a tax of τ on a product induces a reduction in the marginal revenue of paid employees from $P_i \frac{\partial Q}{\partial L_{iP}}$ to $\frac{1}{1+\tau} P_i \frac{\partial Q}{\partial L_{iPE}}$ such that establishments produce at $P_i \frac{\partial Q}{\partial L_{iPE}} > W_{iPE}$, where W_{iPE} denotes firm i's wage rate for paid workers.

Third, in the presence of adjustment costs of paid employees, the s-S-type modelling becomes suitable. In that case, there exists ranges of product demand or technical efficiency shocks such that the plant does not necessarily adjust paid employees every year. Even when paid employment is adjusted, firms do not use first-order conditions to determine employment. Thus, whether the concerned labour input is adjusted or not, the process does not lead to $P_i \frac{\partial Q}{\partial L_{iPE}} = W_{iPE}$.

Fourth, since reallocation growth of paid employment is consistently positive every year, then the manufacturing sector is dominated by firms with either $\Delta ln L_{iPEt} > 0$ and $(\varepsilon_{ik}^v - s_{ik}^v) > 0$ or $\Delta ln L_{iPEt} < 0$ and $(\varepsilon_{ik}^v - s_{ik}^v) < 0$ in Eq.15. Producers of manufactured goods with value-added elasticity with respect to paid labour greater than the revenue share of paid employment for growing incumbent firms contributes positively to APG. Similarly, producers of goods with value-added elasticity less than the revenue share of paid labour for contracting firms contributes positively to APG as well.

Overall, consideration of resource shuffling across plants based on the microfoundations approach produces results similar to those generated by Bailey *et al*. (1992)/Forster *et al*. (2001).This process led us to separate out the reallocation of total labour, paid workers, and working proprietors from total input reallocation. The finding is that, on a year-to-year basis, all input reallocation has a positive impact on APG. More importantly, the component of labour that is widely used by the IMF in country reports for Swaziland; that is, paid employees, is significantly positive every year. It dominates labour reallocation and accounts for 98 percent of all labour shuffled from low to high VMP producers. However, the annual average productivity for primary input reallocation, though still positive, is much lower due to the inclusion of real capital stock. This is due to high capital irreversibility characterizing the manufacturing sector and is likely to constrain entry-exit dynamics while also promoting coexistence of both efficient and inefficient plants.

In the case of net-entry, mergers and acquisitions involving two large firms had a large effect on APG due to the high level of concentration in most industrial sectors. That is, in 1998 a division of a large company was taken over by another firm in the same sector but this was recorded as entry of a new

firm. In the following year the acquiring firm took over the rest of the company and engaged in extensive retrenchments which raised labour productivity in this sector. This behaviour accounted for approximately 265 percent productivity growth in these two years.

3.7   Summary and Conclusion

This chapter investigates primary input trends, aggregate productivity and factor-intensities in Swazi manufacturing firms over a period of trade liberalisation in the Southern African Customs' Union. It begins with descriptive analyses and then investigates the drivers of aggregate productivity growth over time and across industries. A cross-country comparison of drivers of aggregate labour productivity growth with those of the Swazi manufacturing sector is also undertaken. The chapter then deepens the analysis to focus on Swaziland by decomposing aggregate labour productivity growth over time using traditional methods and also relying on Petrin and Levinsohn (2012) as applied by Nishida *et al*. (2014). It concludes with an analysis of seemingly outlying aggregate labour productivity growth in 1998 and 1999 to determine the characteristics of entrants associated with it.

The descriptive evidence shows a decline in both aggregate labour and capital productivities and an increase in the capital−labour ratio. It also shows a leftward distribution of ALP and increasing heaviness of both tails. There are three potential explanations for this. First, firms shed more labour relative to capital due to capital irreversibility and to South African companies shifting production back to South Africa as a response to the lifting of economic sanctions whilst keeping Swazi plants in operation to cover their variable costs. Second, lower productivity firms are growing faster relative to higher productivity plants. Third, there is entry of lower ALP firms.

An in-depth analysis using the conventional approach found that the ALP growth is driven largely by net entry, then by cross-firm market share shift and negatively by within-firm technical change. This result is robust to controlling for confounding effects of plant turnover in the Baily *et al*. (1992) method. Using the Petrin and Levinsohn (2012) approach also produces the same order of importance for APG components. That is, the net-entry contribution explains most of the changes in APG followed by input reallocation, while technical efficiency remains negative per year.

However, the most interesting case is the combined input reallocation reflecting cross-plant movements. The average reallocation of the input bundle from low- to high-productivity incumbent plants is 0.15 percent per year. However, isolating the average annual rate of labour reallocation from the contribution of all inputs put together produces 3.25 percent. Furthermore, paid employment shows positive growth in every year and accounts for an average of about 98 percent of all labour reallocated per year. These results are robust to 'single-deflation' by the manufacturing value-added deflator and 'double-deflation' by consumer price index. Furthermore, the annual average of net-entry

contribution to APG is 58.01 percent and is mainly accounted for by the dramatic increase of APG in 1998 and 1999 due to firm entry.

Finally, the analysis reveals that individual contributions by the extensive and intensive margins of resource reallocation to APG decisively *dominate* technical efficiency in the manufacturing sector in Swaziland. Firms were not investing more in improving production efficiency through innovation and adoption of new technologies than they were moving labour to higher activity producers. This conclusion remained robust regardless of the deflation procedure used in the estimation of the real value-added production function. The novelty of our results lies in the use of micro-foundations to define aggregate productivity growth.

Our future research will focus on separating the contribution of each factor and intermediate input to APG. Given that the APG framework nests many situations around the development and introduction of new goods, this enquiry should also estimate fixed costs and the "gap" terms in Eq. 15 to further understand the productivity dynamics during a period of market reforms. Petrin *et al*. (2011) estimate the impact of primary and intermediate inputs on productivity growth and estimate the orders of magnitude and potential volatility of input gaps. Petrin and Sivadasan (2013) use input gaps to estimate output losses due to allocative inefficiency.

APPENDICES

Appendix A3.1: Manufacturing and Survey ALP (1994−2003)
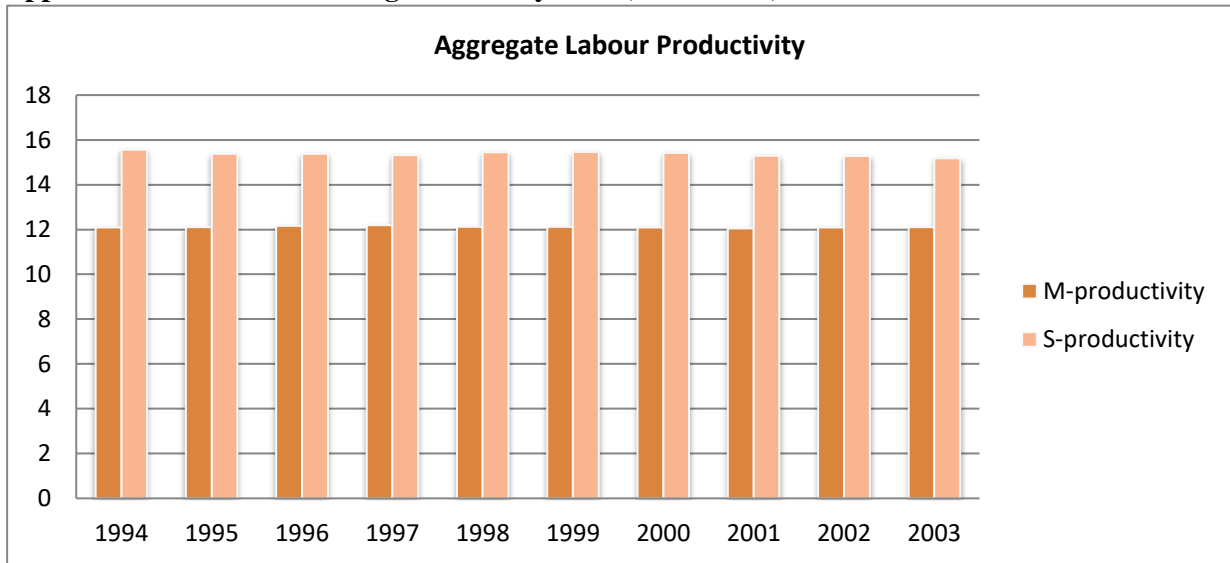
Note: S-productivity denotes ALP measured by the natural logarithm of real value added/labour ratio calculated from survey data and the equivalent M-productivity calculated from real value added sourced from the World Bank Indicators and paid labour sourced from IMF Country Reports for Swaziland (1999, 2000, 2003, and 2008).

Appendix A3.1: ALP Distribution for Selected Years (1994, 1997, 2000, 2003)
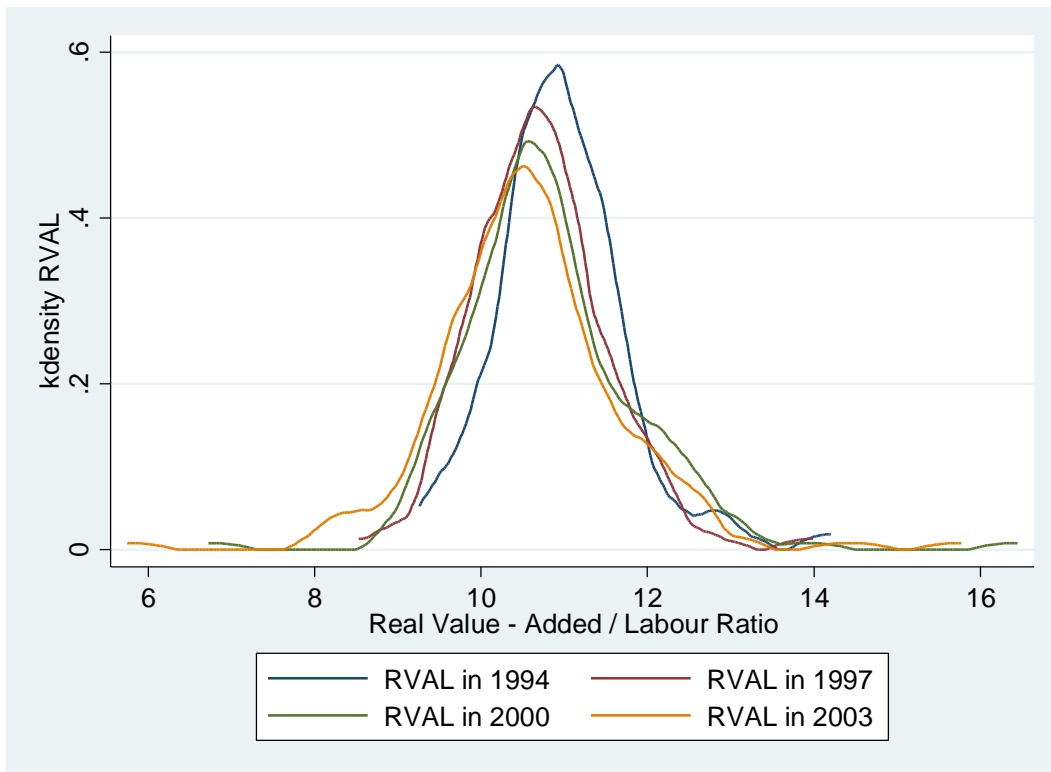
Note: Single deflation of the ratio of real value added to aggregate annual employment.

Appendix A3.2: Evolution of the 25th Percentile of ALP by Industry (1994-2003)

isic2	EVOLUTION OF FIRST QUARTILE ALP BY INDUSTRY									
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Food (15)	1.00	0.70	0.83	0.88	0.91	0.81	0.73	0.74	0.68	0.52
Textile (17)	1.00	0.61	0.58	0.73	1.37	0.72	0.47	0.65	0.17	0.34
Apparel (18)	1.00	-16.15	16.46	10.99	13.17	9.77	11.50	-23.31	7.47	3.95
Wood (20)	1.00	0.66	0.68	0.64	0.61	0.55	0.81	0.80	0.58	0.25
Pulp & Paper (21)	1.00	1.27	1.55	1.51	1.60	1.63	1.49	0.73	1.55	-0.11
Printing & Publishing (22)	1.00	0.93	0.92	0.83	0.79	0.86	0.93	0.69	0.74	0.62
Chemicals (24)	1.00	1.11	0.82	1.08	0.80	0.99	1.03	0.80	0.89	0.95
Rubber (25)	1.00	1.02	0.93	0.77	0.87	0.81	0.90	0.95	0.70	0.51
Non-Metallic Minerals (26)	1.00	0.96	0.61	0.40	0.43	0.47	0.62	0.22	0.41	0.59
Basic Metals (27)	1.00	-0.05	3.13	3.25	3.13	3.16	1.49	1.18	0.17	2.00
Fabricated Metal (28)	1.00	0.84	1.23	0.92	1.10	1.30	0.94	0.80	0.74	1.02
Furniture (29)	1.00	0.94	0.90	1.10	1.07	0.96	0.86	1.02	1.07	0.27
Other Manufacturing (36)	1.00	1.75	0.95	1.03	0.85	1.24	1.20	0.59	0.82	1.02
Sector Mean	**1.00**	**-0.42**	**2.28**	**1.86**	**2.05**	**1.79**	**1.77**	**-1.09**	**1.23**	**0.92**
Sector Median	**1.00**	**0.93**	**0.92**	**0.92**	**0.91**	**0.96**	**0.93**	**0.74**	**0.74**	**0.59**
Std Dev (σ_{ALP})	**0.00**	**4.75**	**4.31**	**2.83**	**3.41**	**2.50**	**2.94**	**6.68**	**1.91**	**1.05**

Source: Author's calculations.

Appendix A3.3: Evolution of the 75th Percentile of ALP by Industry (1994-2003)

isic2	EVOLUTION OF THIRD QUARTILE ALP BY INDUSTRY									
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Food (15)	1.00	0.88	0.94	0.94	1.05	1.06	1.07	1.00	0.95	0.92
Textile (17)	1.00	0.63	0.81	0.97	0.94	1.13	0.66	0.83	0.50	0.77
Apparel (18)	1.00	1.11	1.12	0.70	1.38	0.92	0.73	0.62	0.64	1.54
Wood (20)	1.00	0.84	0.66	0.79	0.99	1.17	1.06	0.73	0.77	0.69
Pulp & Paper (21)	1.00	1.00	1.01	1.25	1.20	1.22	1.42	1.34	1.14	1.09
Printing & Publishing (22)	1.00	0.82	0.87	0.87	1.00	1.03	1.08	1.08	0.97	1.01
Chemicals (24)	1.00	0.95	0.95	1.29	0.96	1.04	1.13	1.20	1.33	0.95
Rubber (25)	1.00	0.85	0.93	0.70	0.87	0.88	0.76	0.65	0.62	0.58
Non-Metallic Minerals (26)	1.00	0.83	0.80	0.65	0.84	0.91	0.82	0.87	0.98	0.90
Basic Metals (27)	1.00	0.99	0.70	0.73	0.70	0.71	0.55	0.53	1.00	0.96
Fabricated Metal (28)	1.00	0.84	0.92	0.75	0.76	0.89	0.83	0.84	0.90	0.83
Furniture (29)	1.00	0.91	0.87	0.78	0.78	0.78	0.74	0.82	0.87	1.34
Other Manufacturing (36)	1.00	0.86	0.84	0.54	0.71	1.12	0.88	0.85	0.73	0.74
Sector Mean	**1.00**	**0.89**	**0.88**	**0.84**	**0.94**	**0.99**	**0.90**	**0.87**	**0.88**	**0.95**
Sector Median	**1.00**	**0.86**	**0.87**	**0.78**	**0.94**	**1.03**	**0.83**	**0.84**	**0.90**	**0.92**
Std Dev (σ_{ALP})	**0.00**	**0.11**	**0.12**	**0.22**	**0.19**	**0.15**	**0.24**	**0.23**	**0.23**	**0.26**

Source: Author's calculations.

Appendix A3.4: Estimation of the Wooldridge-Petrin-Levinsohn Production Function

This Appendix relies on Petrin, Poi and Levinsohn (2004), Galuščák and Lizal (2011) and Wooldridge (2009). The value-added function is specified as in Levinsohn and Petrin (2003):

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + v_{it} , \qquad (1.1)$$

where all variables are expressed in the natural logarithm. β_0 is a constant term, the coefficients (β_l, β_k) are output elasticities with respect to labour and capital, in that order. The unobserved productivity is ω_{it} and v_{it} is a sequence of shocks that is assumed to be conditionally mean independent (CMI) of current and past inputs.

The demand for intermediate inputs is assumed to be a function of capital and the unobserved productivity

$$m_{it} = f(k_{it}, \omega_{it}). \qquad (1.2)$$

Levinsohn and Petrin (2003) demonstrate the monotonicity property of the demand function for intermediates under mild assumptions which allow for the inversion of Eq. 1.2 as

$$\omega_{it} = g(k_{it}, m_{it}) \qquad (1.3)$$

and productivity adjusts according to a Markov process as

$$\omega_{it} = E(\omega_{it}|\omega_{it-1}) + \xi_{it} \qquad (1.4)$$

where ξ_{it} is productivity innovation.

Then, (1.1) can be expressed as either

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + g(k_{it}, m_{it}) + v_{it} \qquad (1.5)$$

or

$$y_{it} = \beta_l l_{it} + \phi(k_{it}, m_{it}) + v_{it} \qquad (1.6)$$

where

$$E(v_{it}|l_{it}, k_{it}, m_{it}) = 0 \qquad (1.7)$$

and

$$\phi(k_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + g(k_{it}, m_{it}). \qquad (1.8)$$

To complete the first stage, the function ϕ in Eq. 1.6 is approximated with a third-degree polynomial in k_{it} and m_{it}, and β_l is estimated using O.L.S.

The final stage sets out to identify β_k. First, the values of Eq. 1.6 are estimated as

$$\hat{\phi}_{it} = \hat{y}_{it} - \hat{\beta}_l l_{it}. \qquad (1.9)$$

Then, using a potential estimate for β_k, say β_k^*, it is possible to estimate the productivity series as

$$\widehat{\omega}_{it} = \hat{\phi}_{it} - \beta_k^* k_{it}. \qquad (1.20)$$

In terms of Levinson and Petrin (2003), a consistent nonparametric approximation to $E(\omega_{it}|\omega_{it-1})$ is given by the predicted values from the nonlinear regression

$$E(\widehat{\omega_{it}|\omega_{it-1}}) = \widehat{\omega}_{it} = \gamma_0 + \gamma_1 \omega_{it} + \gamma_2 \omega_{it}^2 + \gamma_3 \omega_{it}^3 + \vartheta_{it} \qquad (1.21)$$

Thus, given $E(\omega_{it}|\omega_{it-1})$, $\hat{\beta}_l$ and β_k^*, the estimate of β_k solves the minimization of the squared regression residuals

$$\min_{\beta_k^*} \sum_i \left(\hat{y}_{it} - \hat{\beta}_l l_{it} - \widehat{\beta_k^*} k_{it} - E(\omega_{it} \widehat{|\omega_{it-1}}) \right)^2. \tag{1.22}$$

This procedure closes with a bootstrap based on random sampling from observations to construct standard errors of the capital and labour coefficient estimates as in Horowitz (2001).

In stark contrast to the two-step approach, Wooldridge (2009) proposes to simultaneously estimate the capital and labour coefficients by assuming CMI of the i.i.d. error term with respect to current and past values of l_{it}, k_{it}, m_{it} .

CMI Assumption I:

$E(v_{it} | l_{it}, k_{it}, m_{it} , l_{it-1}, k_{it-1}, m_{it-1} , \dots , l_{1t}, k_{1t}, m_{1t}) = 0$. This means the error term is conditional mean independent of, or uncorrelated with, the present and past production inputs. ∎

Wooldridge (2009) restricts the dynamics of the unobserved productivity shocks and writes

$$E(\omega_{it} | l_{it}, k_{it}, m_{it} , l_{it-1}, k_{it-1}, m_{it-1} , \dots , l_{1t}, k_{1t}, m_{1t})$$
$$= E(\omega_{it} | \omega_{it-1})$$
$$= j(g(k_{it-1}, m_{it-1}) \tag{1.23}$$

where $\omega_{it-1} = g(k_{it-1}, m_{it-1})$ and the productivity innovation a_{it} can be written as

$$\omega_{it} = j(\omega_{it-1}) + a_{it}. \tag{1.24}$$

The innovation in Eq 1.24 may reflect heterogeneity and persistence in firm-level managerial ability, labour quality, etc.; see Gebreeyesus (2008).

CMI Assumption II:

$E(a_{it} | k_{it}, l_{it-1}, k_{it-1}, m_{it-1} , \dots , l_{1t}, k_{1t}, m_{1t}) = 0$. Given the quasi-fixed nature of capital in firms due to irreversibility (see, for example, Caballero and Engel, 1999 and Bertola and Caballero, 1994), the productivity innovation a_{it} is uncorrelated with the state variable k_{it} and all past inputs and their functions, but correlated with l_{it} and m_{it} .∎

Substitution of Eq. 1.23 and Eq. 1.24 into Eq. 1.1 yields

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + j(g(k_{it-1}, m_{it-1})) + u_{it} \tag{1.25}$$

where $u_{it} = a_{it} + v_{it}$. Notably, the arguments in the $j(g(k_{it-1}, m_{it-1}))$ function are now lagged capital and intermediate inputs which can be approximated with low-order polynomials as in Levinsohn and Petrin (2003).

CMI Assumption III:

$E\left(u_{it} | k_{it}, l_{it-1}, k_{it-1}, m_{it-1} , \dots , l_{1t}, k_{1t}, m_{1t,} \right) = 0$. The error u_{it} is conditional mean independent of current capital and past values of all production inputs. In the presence of the productivity innovation in u_{it}, this condition is identical to Conditional Mean Independence Assumption II above. ∎

Therefore, Eq 1.1 becomes

$$y_{it} = \varphi_0^* + \beta_l l_{it} + \beta_k k_{it} + g(k_{it-1}, m_{it-1}) + u_{it} \tag{1.26}$$

or

$$y_{it} = \varphi_0^* + \beta_l l_{it} + \beta_k k_{it} + \sum_p^3 \sum_q^{3-p} \hat{\delta}_{pq} k_{it-1}^p m_{it-1}^q + u_{it}. \blacksquare\blacksquare\blacksquare \qquad (1.27)$$

Appendix A3.5: Proportion of Non-Zero Input Observations

Industry (ISIC)	Investment	Material	Energy
Food and Food Products(15)	45.39	100.00	94.09
Textile (17)	30.51	100.00	99.44
Apparel (18)	20.31	100.00	100.00
Wood and Wood Products (20)	35.51	100.00	90.65
Paper and Paper Products (21)	61.82	100.00	89.09
Printing, Publishing (22)	23.12	100.00	96.48
Chemicals and Chemical Products (24)	26.36	100.00	90.70
Rubber and Plastic Products (25)	49.09	100.00	98.18
Other non-metallic Minerals (26)	29.45	100.00	93.25
Basic Metals (27)	9.68	100.00	100.00
Fabricated Metal Products (28)	33.16	100.00	91.98
Machinery and Equipment (29)	46.00	100.00	100.00
Furniture and Other Manufacturing (36)	32.32	100.00	97.98
Average	34.06	100.00	95.53

Source: Author's calculations from Data Compiled by the CSO

Appendix A3.6: Aggregate Multifactor Productivity Growth Rate, Swaziland Manufacturing 1994–2003: LP APG Decomposition, Manufacturing Value-added Index Single-deflator.

Year	Value-Added Growth	APG (0)	APG Decomposition: (0)= (1) + (2) + (3)					Net Entry (3)
			Technical Efficiency (1)	Reallocation				
				Total Reallocation (2)	Labour Reallocation	Working Proprietors Reallocation	Paid Employees Reallocation	
1995	7.76	2.51	0.91	-9.65	1.39	-1.48	2.87	11.25
1996	23.10	16.03	3.17	-7.89	0.92	-0.16	1.08	20.74
1997	-44.35	-34.12	-3.08	18.52	9.31	-0.06	9.37	-49.56
1998	265.55	240.84	1.67	-1.74	0.69	0.02	0.67	240.91
1999	275.57	261.04	1.23	9.22	2.94	-0.02	2.95	250.59
2000	-16.28	-16.07	-14.97	-5.48	0.42	-0.14	0.56	4.38
2001	37.42	34.85	8.71	20.43	0.42	0.43	-0.01	5.72
2002	-20.74	-21.81	-3.04	-29.53	-1.15	-1.21	0.06	10.76
2003	-36.71	-33.05	-22.31	2.69	8.35	-1.62	9.97	-13.43
Mean	**54.59**	**50.02**	**-2.77**	**-0.38**	**2.33**	**-0.42**	**2.75**	**53.48**
Median	**7.76**	**2.51**	**0.91**	**-1.74**	**0.92**	**-0.14**	**1.08**	**10.76**
Std Dev	**125.32**	**116.24**	**9.66**	**15.46**	**3.70**	**0.75**	**3.90**	**110.93**

Note: As in Nishida *et al.* (2014), numbers are percentage growth rates. The plant-level multifactor productivity is calculated by using production function parameters that vary across 2-digit ISIC. We obtain the estimates by using Wooldridge (2009). APG represents the aggregate productivity growth with entry and exit, which is defined as aggregate change in final demand *minus* aggregate expenditure in inputs, holding input constant. We use value-added share (Domar) for weights. APG is decomposed into four components: (1) technical efficiency, (2) reallocation, and (3) net entry term, using Eq. 17 in text.

Appendix A3.8: Aggregate Multifactor Productivity Growth Rate, Swaziland Manufacturing 1994–2003: LP APG Decomposition, Consumer Price Index Double-deflator.

| Year | Value-Added Growth | APG (0) | Technical Efficiency (1) | APG Decomposition: (0)= (1) + (2) + (3) | | | | Net Entry (3) |
| | | | | Reallocation | | | | |
				Total Reallocation (2)	Labour Reallocation	Working Proprietors Reallocation8	Paid Employees Reallocation	
1995	27.96	19.22	7.23	-0.95	2.53	-0.89	3.41	12.94
1996	21.85	11.94	1.63	-7.17	1.40	-0.30	1.71	17.48
1997	-41.95	-27.47	-1.81	20.30	10.07	-0.05	10.11	-45.96
1998	268.72	235.19	1.93	-1.48	-0.04	0.01	-0.05	234.74
1999	270.96	251.41	0.97	8.69	3.01	-0.01	3.01	241.75
2000	-17.72	-17.44	-16.14	-6.01	0.61	-0.18	0.79	4.72
2001	42.77	39.22	9.56	23.23	0.53	0.44	0.09	6.44
2002	-23.53	-24.96	-4.76	-31.25	-1.23	-1.28	0.05	11.06
2003	-36.51	-31.45	-22.93	3.59	9.67	-0.43	10.10	-12.10
Mean	**56.95**	**50.63**	**-2.43**	**0.99**	**2.65**	**-0.27**	**2.92**	**52.34**
Median	**21.85**	**11.94**	**0.97**	**-0.95**	**1.40**	**-0.18**	**1.71**	**11.06**
Std Dev	**124.24**	**111.90**	**10.59**	**16.20**	**4.12**	**0.52**	**4.09**	**107.13**

Note: As in Nishida *et al.* (2014), numbers are percentage growth rates. The plant-level multifactor productivity is calculated by using production function parameters that vary across 2-digit ISIC. We obtain the estimates by using Wooldridge (2009). APG represents the aggregate productivity growth with entry and exit, which is defined as aggregate change in final demand *minus* aggregate expenditure in inputs, holding input constant. We use value-added share (Domar) for weights. APG is decomposed into four components: (1) technical efficiency, (2) reallocation, and (3) net entry term, using Eq. 17 in text.

CHAPTER 4: Investment Dynamics, Unobserved Heterogeneity and Endogenous Investment Switching Regime in Manufacturing

4.1 Introduction

The purpose of this chapter 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. This is an interesting period in its own right because of observed micro churning dynamics and industrial reorganization induced by the trade reforms. It is also interesting because it uses an unbalanced firm-level panel data set that has never been used before to identify determinants of investment decisions in the manufacturing sector. The achievement of this goal is important for both policymakers and economic practitioners working in the field who face the task of investigating investment patterns in plant, machinery and equipment in the presence of a high incidence of zero investments.

Typically, the model 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).[42] An alternative candidate is the ratio of the firm's stock market value to its capital replacement cost; that is, Tobin's average q.

---

[42] 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.

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 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).

The purpose of this chapter is to estimate a structural model of investment to determine 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[43].

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

---

[43] 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.

(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. 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.[44] 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

---

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

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, 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 chapter is organized as follows: The next section describes the panel dataset and performs descriptive analyses of investment rates for the manufacturing sector. In Section 4.3, the shape of the empirical hazard and fixed adjustment costs are investigated while Section 4.4 discusses econometric models and estimators for the structural model of investment, emphasising the General Method of Moments' approach. Section 4.5 considers alternative methods based on nonlinear dynamic random effects techniques in the estimation of binary structural models. The characteristics of endogenous investment switching by firms in manufacturing are addressed in Section 4.6. Empirical results are presented in Sections 4.7-4.8 and Section 4.9 concludes the analysis.

4.2 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. 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 investment variable response, expenditure in and sales of PME are reported either with missing values or with real numbers. In structural equations, 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 logarithms. 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.[45]

Table 4.1 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 of a typical proxy of marginal q is approximately $1\frac{1}{3}$ times higher than that of the investment rate. The longitudinal investment rate of variation measured by the standard deviation is relatively low at 0.29, with an average investment rate of 0.24 and $(\text{Min}, \text{Max}) = (-0.83, \ 1.58)$.[46] 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 one of the proxies to study the behaviour of investment rates is likely to suffice.

There are at least two explanations for the patterns observed in Table 4.1. 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 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.

---

[45] 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.

[46] 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.

Table 4.1: Summary Moments of Key Variables

Statistics	Key Variables					Proxies of Marginal q		
	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.

Table 4.2 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 4.2: The Correlation Matrix of the Main Variables

	$\dfrac{I_t}{K_{t-1}}$	$\dfrac{I_{t-1}}{K_{t-2}}$	$\dfrac{I_{t-2}}{K_{t-3}}$	$\dfrac{I_{t-3}}{K_{t-4}}$	$\dfrac{I_{t-4}}{K_{t-5}}$	K_t^{-1}	Emp_{t-1}	$\dfrac{S_t}{K_{t-1}}$	$\dfrac{CF_t}{K_{t-1}}$	$\dfrac{P_t}{K_{t-1}}$
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 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 4.2 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[47].

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 4.1. 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 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 4.1: Investment Rate Relationship with the Sales/Capital Ratio
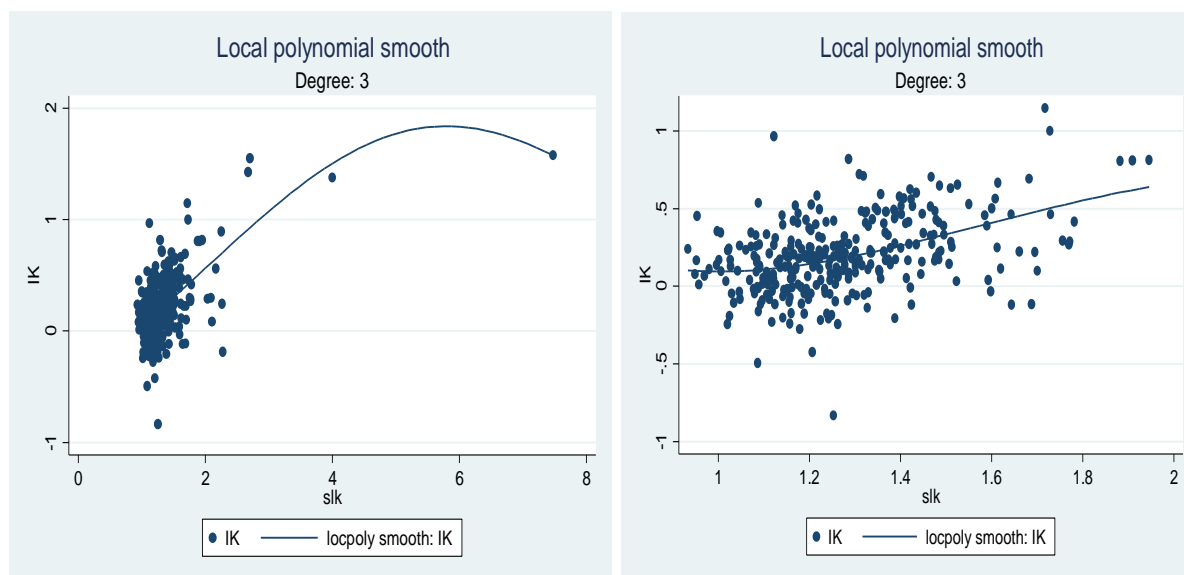
As in Cooper, Haltiwanger and Power (1999) and Cooper and Haltiwanger (2006), the rest of this chapter defines net investment in terms of real gross expenditure (EXP_{it}) on PME and real sales $(SALES_{it})$ for firm i at time t for the class of capital goods concerned. One striking feature of the expenditure series is that it isolates the cost of maintenance and repairs, permitting a sharper investigation of non-smoothness of (dis)investments. Hence

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

and

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

---

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

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 4.2, 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.[48] 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 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 4.2a. 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 and high kurtosis in the distribution of investment rates is indicative of the presence of nonconvexities in the capital adjustment technologies and a presence. Fat tails to the right suggest the presence of a large fraction of large capital adjustments.

Figure 4.2: Distribution of Investment Rates of PME.
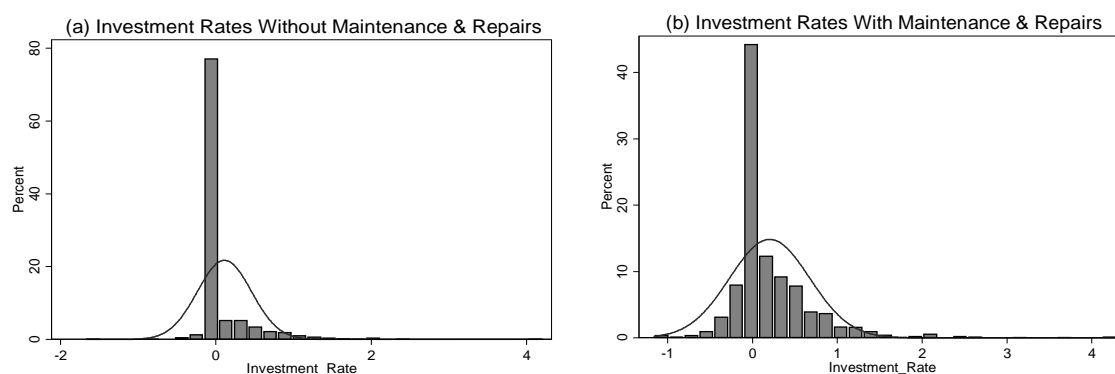
In summary, the analysis thus far provides several lessons. It reveals that there is low propensity to invest in capital goods and that the observed heterogeneity in the rate of investment is just as low. As is typical in the literature, investment inactivity dominates the distribution of investment 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 suggest the presence on nonconvexities in the capital adjustment costs. Firm-level investment behaviour, including financially unconstrained firms, is also consistent with increased focus on M&R. These patterns imply the

---

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

presence of low costs of capital adjustment in manufacturing and a high rate of obsolescence in machinery and equipment needed for use in production.

4.3 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, Collier, Dercon, Fafchamps, Gauthier, Gunning, Oostendorp, Pattillo, Soderbom and Teal (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 $\frac{I_{it}}{K_{it-1}} = 0$. [49]

---

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

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 4.3, firm scale-independence means the null hypothesis H_0: $\text{Survivor}_{\text{Large}} = \text{Survivor}_{\text{Small}}$ is not true, where large firms employ more than 50 workers. The Peto-Peto-Prentice test does not support the null hypothesis at the 1percent level.[50] 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.[51] 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 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 $=1$).

---

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

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

Figure 4.3: Kaplan-Meier Survival and Hazard Estimates of Investment
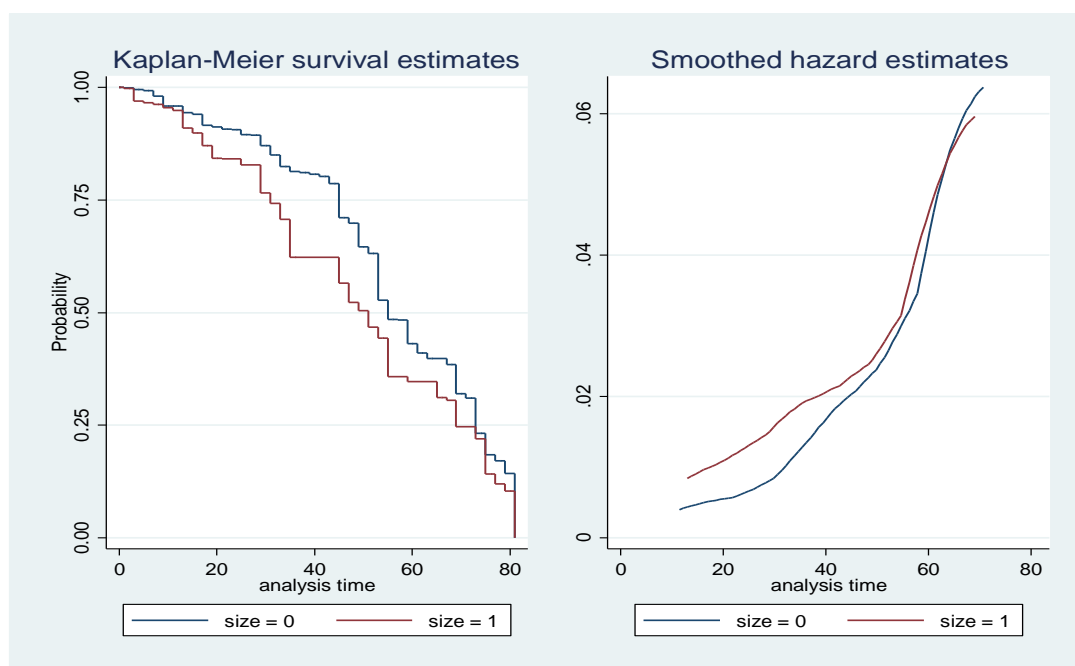
The right panel in Figure 4.3 plots the empirical hazard expressed in Eq. 4 against the time since the last investment. It shows that the probability of having a lumpy investment episode is scale-dependent, where larger firms have a relatively higher probability of an investment spike compared to smaller 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.[52]

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 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 the within-firm effects rather than between-firm effects in PME investments.

---

[52] 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.

4.4 Econometric Models and Estimators

The purpose of this section is to present models and associated estimators that are useful in determining the probability of investing in durable capital goods. It 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.

For continuous responses, the long tradition of GMM approaches in estimating dynamic panel data models dominates empirical research in continuous response environments, see Arellano and Bond (1991) and Blundell and Bond (1998). In the case of binary response models of investment, a distinction between true state dependence and unobserved heterogeneity is normally achieved through dynamic modelling that includes a lagged response and a random intercept. The multilevel framework of analysis can also 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). In both continuous and binary models of investment, the assumption is that firms do not sort according to whether a firm belongs in a high or low investment regime. The final model therefore closes this gap by distinguishing between firms in high and low investment regimes, see Lee and Frost (1978), Maddala (1983) and Lokshin and Sajaia (2004).

4.4.1   The GMM Approach

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

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

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, small T DPD structure is assumed. The measure of state dependence $|\boldsymbol{\alpha}| < 1$[53] 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, $\boldsymbol{x}_{it}$

$$y_{it} = \boldsymbol{\alpha} y_{it-1} + u_{it} \qquad\qquad (6)$$

---

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

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

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

The α-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.

4.4.1.1 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

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

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

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

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

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

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 $\mathbf{Z}_d = (Z_{d1}', Z_{d2}', \dots, Z_{dN}')'$ and W_N is a two-step weighting matrix assuring validity of efficiency properties for the GMM estimator. The matrix is defined as

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

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 $\Delta \hat{u}_i = \Delta y_{it} - \hat{\alpha} \Delta y_{it-1}$ 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).

4.4.1.2 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

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

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

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

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

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

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

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

100

$$E(Z'_{si}p_i) = 0$$

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

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

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

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

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

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).[54]

4.4.2   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

---

[54] The problem of high autoregressive parameter; $\hat{\alpha} \to 1$ and $(\sigma_{\eta i}/\sigma_{vi}) \to \infty$, 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), and Youssef and Abonazel (2015).

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 $c_{it} = \sqrt{\frac{T_{it}}{(T_{it}+1)}}$.[55] 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

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

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

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

$$M_{\perp} = \left\{\begin{array}{cccc} \sqrt{\frac{(T-1)}{T}} & \frac{1}{\sqrt{T(T-1)}} & \frac{1}{\sqrt{T(T-1)}} & ... \\ & \sqrt{\frac{(T-2)}{(T-1)}} & \frac{1}{\sqrt{(T-1)(T-2)}} & ... \\ & & \sqrt{\frac{(T-3)}{(T-2)}} & ... \end{array}\right\}$$

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. $M_{\perp}M_{\perp}' = I$. This is an expression portraying the assumption of homoscedasticity carried out in Arellano and Bond (1991).

---

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

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), 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).

4.5 Nonlinear Dynamic Random-Effects Models and Estimators

4.5.1 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

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

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

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 $\mathbf{z}_j$ and time-varying $\mathbf{x}_{ij}$ 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, $h\big(p_{ij}\big) = \Phi^{-1}\big(p_{ij}\big)$, where $p_{ij} = \text{pr}\big(y_{ij} = 1 | y_{i-1,j}, \mathbf{z}_j, \mathbf{x}_{ij}, \zeta_j\big)$. 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

$$\begin{cases} \Pr(y_{ij} = 1|y_{i-1,j}, \mathbf{z}_j, \mathbf{x}_{ij}, \zeta_j) = h^{-1}(\mathbf{z}_j'\boldsymbol{\gamma_z} + \mathbf{x}_{ij}'\boldsymbol{\gamma_x} + \overline{\mathbf{x}}_{.j}'\boldsymbol{\delta_{\bar{x}}} + \boldsymbol{\alpha}y_{i-1,j} + u_j) \\ \qquad \Pr(y_{0j} = 1|\mathbf{z}_j, \mathbf{x}_j, \zeta_j) = h^{-1}(\mathbf{z}_j'\boldsymbol{g_z} + \mathbf{x}_{0j}'\boldsymbol{g_x} + \overline{\mathbf{x}}_{.j}'\lambda_0\boldsymbol{\delta_{\bar{x}}} + \lambda_0 u_j) \end{cases}, \text{i=1, ..., T-1} \qquad (9)$$

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

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

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.

4.5.2   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

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

where $\sum p_j = 1, p_j \geq 0, j = 1, \ldots, k, k$ is the number of points of support, p_j is probability mass point, θ_j is a locator of p_j such that $p_j = \text{prob}(\theta = \theta_j)$. 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

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

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 $g(y_i|X_i, \theta_i)$ is supported. The Heckman-Singer estimator has been found consistent for mixing distributions with a small number of points of support.

4.6 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 the different regimes. The goal here is to estimate the switching regression model specified in Eq. 11

$$\begin{cases} \frac{I_{it}}{K_{it-1}} = X_{it}\beta^{Low} + \varepsilon_{1it} & \text{iff} & Z_{it}\gamma + u_{it} < 0 \\[3mm] \frac{I_{it}}{K_{it-1}} = X_{it}\beta^{High} + \varepsilon_{2it} & \text{iff} & Z_{it}\gamma + u_{it} \geq 0 \end{cases} \qquad (11)$$

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. 16 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. 16 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.

4.7 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 practice 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

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

where the two-way error structure is defined as $u_{it} = \mu_i + \tau_t + \varepsilon_{it}$, for $t = 2, \dots, T$, μ_i and τ_t are the unobservable firm-specific effect and time effects, respectively; while ε_{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 and its lag is included as a proxy for marginal q, while the employment regressor controls for primary input substitution effects. Eq.12 is estimated using the

Generalized Method of Moments (GMM), the random effects approach and the endogenous investment regime switching method.

4.7.1   The GMM Estimates

4.7.1.1 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. However, 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), Blundell, Bond and Windmeijer (2000).

Table 4.3 summarizes the empirical anatomy of section 4.4. 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.[56] 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 4.3: Schema for the Empirical GMM Analysis Using Arellano and Bond (1991) for $\hat{\alpha}_{GMM\,Diff}$ and Blundell and Bond (1998) for $\hat{\alpha}_{SYS}$

Parameter	$\hat{\alpha}_{Diff} = \dfrac{\Delta y'_{-1}\mathbf{Z}_{D,i}W'_N\mathbf{Z}'_{D,i}\Delta y}{\Delta y'_{-1}\mathbf{Z}_{D,i}W'_N\mathbf{Z}'_{D,i}\Delta y_{-1}}$		$\hat{\alpha}_{SYS} = \dfrac{q'_{-1}\mathbf{Z}_s W_{NS}^{-1}Z'_s q}{q'_{-1}\mathbf{Z}_s W_{NS}^{-1}Z'_s q_{-1}}$	
	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 4.3 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.

Table 4.4 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.

---

[56] See definitions in Appendix A4.1.

Theory argues that although marginal q is a sufficient statistic for investment rates, the sales/capital ratio is also a sufficient statistic for investment rates 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.

Table 4.4: GMM Estimation of Investment Rate Dynamics using an Instrument Reduction Technique and the Helmert's Transform[57]

	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_p^2	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:
1. All models include Year Dummies.

The table 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 $\hat{\alpha}_{Diff}$ and $\hat{\alpha}_{SYS}$ are respectively presented. The first-order autoregressive parameter is high; that is, $\hat{\alpha} \to 1$, [and it might also be the case that $(\sigma_{\eta i}/\sigma_{vi}) \to \infty$], implying 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)). This may be a reflection of an imprecisely measured parameter due to high correlation between the sales/capital variable and omitted variables and other factors. This is the natural characteristic of the GMM DIFF estimator while the GMM SYS estimator circumvents this problem. The one-step GMM SYS estimator has the autoregressive parameter $\hat{\alpha} > 1$, rendering the system non-convergent. The two-step GMM estimator

---

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

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.

Moreover, these results are an outcome of instrument proliferation that is controlled for by first collapsing the instrument count and secondly by using instrument collapsing together with truncation of lag depth to $t - 2$. Our results are invariant to either of the choices. Instrument explosion curtailed by both mechanisms reduces proliferation from 79 to 18 instruments for *Difference* GMM and from 100 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 $\widehat{\alpha}_{Diff}$ results, but h(3) has the effect of slightly increasing the size of $\widehat{\alpha}_{SYS}$ 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.

The diagnostic tests are consistent with a persistent investment rate series and render the two-step System GMM our preferred specification. The principal assumption in the System-GMM estimator is $E(Z'_{si}p_i) = 0$, where p_i is a measure of the combined orthogonal firm-specific effects and idiosyncratic disturbances. Thus, the Two−Step System GMM parameter means that the unobserved group effects among firms are *un*correlated with first-order differences in instrumental variables. Put differently, as in Bontempi and Golinelli (2014), the covariance between firm-specific effects and instruments is constant over time.

It is therefore not possible to draw sound conclusions on whether or not there exists true state dependence in Swazi manufacturing investments based on these 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. In Table 4.5, the empirical model is estimated in full with employment as a control variable for primary input substitutability. Although the size of the AR(1) parameter is substantially reduced across all estimators, it remains insignificant. However, only the two-step GMM SYS (COLL) estimator passes all the Arellano-Bond (1991) diagnostic tests while the rest do not. That is, it satisfies the AR(1) and AR(2) conditions as well as the Sargan and Hansen tests of over-identifying restrictions and instrument validity, respectively. All explanatory variables in this estimator are statistically insignificant at standard levels.

Table 4.5: GMM Estimation of Investment Rate Dynamics with the Control Variable using an Instrument Reduction Technique and the Helmert's Transform

Variables	GMM DIFF (COLL)		GMM SYS (COLL)	
	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:
1. All models include Year Dummies.

A strict 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 4.2 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.[58] 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.[59] 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 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.

4.7.1.2 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 in the previous section, we implement the same model but this time using the first-difference deviations transform. In the absence of controls as shown in Table 4.6, the coefficient of the lagged investment rate variable increases as expected, and 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.

---

[58] 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.

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

Table 4.6: GMM Estimation of Investment Dynamics using the Roodman (2009b) Method of Instrument Reduction with Standard First Difference Deviations Transform without Controls

Variables	GMM DIFF		GMM SYS	
	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 robustness check, Table 4.6 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.[60] 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.

---

[60] 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.

Table 4.7: GMM Estimation of Investment Dynamics using the Roodman (2009b) Method of Instrument Reduction with Standard First Difference Deviations Transform with Controls

Variables	GMM DIFF		GMM SYS	
	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_p^2	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 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.

4.8 Dynamic Random-Effects Estimates

4.8.1 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

$$\text{probit}\left\{\left(\frac{I_{ij}}{K_{i-1,j}} = 1 \middle| \frac{I_{i-1,j}}{K_{i-2,j}}, z_j, x_{ij}, \zeta_j\right)\right\} = \gamma_{z_0} + \alpha\left(\frac{I_{i-1,j}}{K_{i-2,j}}\right) + \gamma_{x_1}\left(\frac{S_{ij}}{k_{i-1,j}}\right) + \gamma_{x_2}(\mathbf{Emp}_{ij}) + \zeta_j$$

where $x_{ij} \in \left[\frac{S_{ij}}{k_{i-1,j}}, \text{Emp}_{i-1,j}\right]$ and there are no time-invariant, z_j, covariates. In this setting, $\frac{I_{ij}}{K_{i-1,j}}$ 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 $i = 0$

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

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

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

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

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

4.8.2 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 4.9.

For each sequence in Table 4.8, a "1" in the ith position denotes an observed positive investment in the ith 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 $\overline{x}_{\cdot j}$ 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)[61]. 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).

---

[61] 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 4.8: Manufacturing Patterns of Missing Values and Investment Participation $\left(y_{ij} = \frac{I_{ij}}{K_{ij-1}}\right)$ 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 4.9: Multilevel Parameter Estimates and Robust Standard Errors for Dynamic Random Effects Probit Models of Investment.

Structural Parameters	Naïve	Estimates for Joint Models		Estimates for Conditional Maximum Likelihood Model	
		Exogenous x_{ij}	Endogenous x_{ij}	Conditional Estimator	NPMLE (*Mass Point Method*)
$\left(\dfrac{I_{i-1,j}}{K_{i-2,j}}\right)$	2.0634**	0.2189	0.5556	0.9143	0.7327
	(0.6817)	(0.7428)	(0.6900)	(0.6566)	(0.6166)
$\left(\dfrac{S_{ij}}{k_{i-1,j}}\right)$	−0.7291	0.2655	0.1574	1.0331	1.2758
	(0.3892)	(0.5596)	(0.4487)	(0.7938)	(0.7857)
$\left(\text{Emp}_{ij}\right)$	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)		
$\left(\dfrac{S_{ij}}{k_{i-1,j}}\right)X$ **Nolag**		1.8706	2.0390		
		(1.1988)	(1.1943)		
$\left(\text{Emp}_{ij}\right)X$ **Nolag**		0.5516**	0.4258*		
		(0.2081)	(0.2084)		
$\left(\dfrac{S_{ij}}{k_{i-1,j}}\right)^{O}$				−2.8371*	−2.5949
				(1.2537)	(1.4484)
$\left(\text{Emp}_{ij}\right)^{O}$				−0.7460	−1.3585
				(0.5493)	(0.7236)
$\left(\dfrac{S_{j}}{k_{1,j}}\right)$				0.7728	0.1701
				(1.0795)	(0.7960)
$\left(\overline{\text{Emp}_{2,j}}\right)$				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:					
$\left(\dfrac{\widetilde{s_{j}}}{k_{1,j}}\right)$			−0.139		
			(0.2169)		
$\left(\widetilde{\text{Emp}}_{2,j}\right)$			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 4.9 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 $\textbf{Emp}_{ij}$, 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.[62] 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

---

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

model insignificantly at all conventional levels.[63] 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 awash with missing values in the investment series as shown in the descriptive analysis of Section 4.2.[64] 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 based on the Gâteaux derivative method. This technique avoids making any assumptions about the

---

[63] 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 $\left(\widetilde{Emp}_{2,j}\right)$, 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.

[64] 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.

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.

4.8.3   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.

In this section we present 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).

This method simultaneously estimates 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).[65] This implies three scenarios:

---

[65] 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.

(1) that if larger firms are more likely to locate in the higher investment regime, the firm-size 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 4.10 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.[66] 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 $H_0: \beta^{High} = \beta^{Low}$ is not supported by the χ^2–distribution of the Wald-test statistic at the 1 percent level.[67] 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.

---

[66] 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.

[67] 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 4.10: FIML Estimation of Endogenous Switching Regression Models: 1994-2003

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

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*** (0.368)	7.396*** (0.2298)	7.715*** (0.3942)	7.203*** (0.4299)	7.208*** (0.333)	7.186*** (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*** (0.4943)	4.553*** (0.4837)	6.058*** (0.497)	3.373 (3.0102)	4.281*** (1.089)	3.904*** (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 (0.1491)	−0.187 (0.1673)	−0.073 (0.3839)	0.785 (3.1523)	1.384 (1.1185)	1.846 (1.301)
	Constant	−0.147 (0.4144)	0.566 (0.3813)	−0.523* (0.2622)	0.479 (0.5853)	0.568 (0.4119)	0.56 (0.4331)
Statistics	$\sigma_{H\varepsilon}$	1.8183 (0.0947)	1.7881 (0.0958)	1.8476 (0.1003)	1.6548 (0.4094)	1.7930 (0.3629)	1.7846 (0.3649)
	$\sigma_{L\varepsilon}$	1.3066 (0.2412)	1.2785 (0.2265)	1.3626 (0.2534)	1.6008 (1.4987)	1.3316 (0.2741)	1.2839 (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: $\qquad\qquad$ χ^2 (1) = 269.67, Prob > χ^2 = 0.0000

$H_0: \beta^{High} = \beta^{Low}$ for Model 4: $\qquad\qquad$ 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.[68] 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.

4.9 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 awash with 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

---

[68] 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.

inactivity for a significant number of establishments in 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 a 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.

4.10   Summary and Conclusion

This chapter 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.

As is typical, investment inactivity dominates the distribution of investment rates, with or without M&R exclusions. Under conditions of new and M&R investment in PME, 44.05 percent of firm-year observations experienced zero investment in the 10-year period. About 14.53 percent and 11.44 percent of these observations experienced a single and two investments in the same period, respectively. Only 36.12 percent of firm-year observations have at least two consecutive, non-missing investment sequences. There is generally a positive correlation between missingness patterns of investment and the likelihood of firms experiencing them.

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.

Such behavioural patterns of investment are consistent with relatively more focus on industrial M&R for machinery and equipment, and much less on new investment in, or replacement of, PME. This implies that the manufacturing sector in Swaziland experienced limited capital adjustment costs of installation and worker re-training. It also experienced marked obsolescence in machinery and equipment due potentially to heightened economic uncertainty.

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

Variable	GMM DIFF−PCA		GMM SYS−PCA	
	One−Step	Two−Step	One−Step	Two−Step
$\dfrac{I_{t-1}}{k_{t-2}}$	−0.014	0.025	0.358	0.358*
	(0.2536)	(0.2817)	(0.1839)	(0.1665)
$\dfrac{s_t}{k_{t-1}}$	−0.187	−0.244	−0.163	−0.162
	(0.1801)	(0.199)	(0.1692)	(0.1584)
$\dfrac{s_{t-1}}{k_{t-2}}$	−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_p^2	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.

CHAPTER 5: Conclusion

This dissertation has two but related purposes. *First*, it exploits a unique firm-level panel dataset of manufacturing firms in Swaziland to extract evidence on micro activities that culminate in macro outcomes during a very interesting period in the history of the Customs Union. *Second*, it answers the following very specific questions:

a) What is the general nature of structural change in the Swazi economy? How has firm-size distribution in the manufacturing sector evolved? Is the popular belief about the job creating prowess of small firms a valid proposition for the manufacturing sector in Swaziland?

b) What impact does firm-level technical efficiency and primary input reallocation across firms have on aggregate productivity growth in manufacturing? As an auxiliary question, how much impact does firm turnover have on APG in the sector?

c) What are the characteristic patterns of industrial investment in plant, machinery and equipment in Swaziland? What effects do state dependence and unobserved heterogeneity have on investment decisions in manufacturing? Is a structural investment model best explained in terms of an investment regime switching model in the manufacturing sector? How can the cost of exercising the investment option to wait be measured in an economic environment replete with uncertainty?

The manufacturing sector in many African economies is characterized by positive growth driven mostly by structural change since the last decade and a half prior to the 21st Century. Reallocation of labour inputs across sectors contributes relatively more than within-firm technical change to aggregate productivity growth. As a result, some leading development economists have dubbed this an 'African Growth Miracle'. It replaces the traditional pessimism of growth prospects with notions of expanded Chinese investment and positive commodity price movements. However, this over-dependence on the external environment, the low levels of productivity and constrained private sector investment in globally competitive industries are conducive to an unsustainable growth path for African economies.

An analysis of structural change focussing on Swaziland showed a persistent weakening of the manufacturing sector in terms of its share of economic activity and employment relative to the services sector. The manufacturing sector's share of GDP trended downwards while the agricultural share of output was mostly fixed at the same level throughout the period of analysis. During the same period, the services sector's share only marginally trended upwards. Therefore, the size of the manufacturing sector in Swaziland, the lack of robust industrialization and the limited diversification into globally competitive industrial investments are potential constraints to structural change. In the large and growing literature, the observation is that economic development in Africa is not likely to come from the manufacturing sector, but rather from either agriculture or services. However, given Swaziland's level of development, economic development driven by the services sector would

123

constitute premature de-industrialization that might render the country's economic growth trajectory unsustainable at best or divergent at worst.

At a micro level, the character of firm-level data was evaluated to establish its quality in relation to published macro data. The annual census data for the manufacturing sector in Swaziland closely resembles similar datasets collected by other statistical agencies. In order to analyse the panel dataset directly, the entry-exit dynamics were measured on the basis of a plant's identity code appearing for the first time rather than on firm registration or the last date of existence in the database, respectively. On the whole, the dataset was at least of as good a quality as any other compiled by a government statistical agency.

Analyses of the data by two-digit ISIC industry show the sector's overdependence on a few primary commodities for export to preferential markets. This exposes producers, upstream suppliers of inputs, and downstream customers to the potential risk of preferential treatment erosion. For example, a loss of market access for sugar in the EU and U.S. would cause the sugar industry to trade in the volatile world market where sugar prices are generally depressed. Export revenue would decline significantly forcing sugar producers to scale down operations. Likewise, upstream sugarcane farmers would receive reduced revenue such that the scale of production would also need scaling down. Again, downstream manufacturers of soft drink concentrates and other users of sugar would have inadequate supplies of this critical input and may have to import it and incur significant transport costs. The effect on the whole value chain would be a loss in revenues and employment.

The notion that the distribution of firm size in developing African countries is characterized by a bimodal distribution with a missing middle is investigated graphically and statistically in the industrial sector. The annual firm-size lognormal distributions shifted towards the left demonstrating a general economic deterioration during the 10-year period. This distribution evolves overtime to a state where it initially declines rapidly after its modal level, then slows down as if to form a 'dip' before accelerating again. The pattern of distributional change does not form two modes at any one year. Using a statistical approach, we regress the average product of primary inputs on firm size. The dip-test statistic for both proxies of the marginal product of inputs strongly rejects the presence of a missing middle in the ten annual cross-sections of the data. Thus, the missing 'missing middle' found in Hsieh and Olken (2014) for the cases of India, Indonesia and Mexico is confirmed for the case of Swaziland.

In the study of job flows, the manufacturing sector produces results that are consistent with findings in other countries. The simultaneity of job creation and destruction features throughout the period of analysis for each establishment size category. However, job destruction significantly dominates job creation in all plant sizes. Looking at longitudinal patterns in detail, small firms destroyed more jobs

than they created and large firms created more jobs than small plants. While small plants create an annual average of 0.9 percent jobs, large plants create 3.5 times more new jobs every year. Job destruction by large firms shows sharp volatility throughout the period of analysis and it also exceeds the job destruction by small plants. Although small firms destroy an annual average of 3.59 percent jobs, large firms destroy approximately as much as 1 in 9 manufacturing jobs every year. Thus, the manufacturing sector was destroying more jobs than creating them as in Kerr *et al*. (2013) for South Africa's case.

Interesting results are produced when job turnover is linked to productivity growth. The longitudinal productivity effect was on average as low as −5.93 percent during the decade of economic reforms in the Customs Union. This means that productivity growth within incumbent firms had a negative impact on the overall productivity growth. Since this average metric is higher than the median, it suggests that the productivity distribution is left heavy-tailed with volatility estimated at 1.7. This accords well with literature that there is prevalence of firm-level inefficiency and limited structural transformation in developing country manufacturing sectors. Labour reallocation is on average 4.63 percent and positively skewed with volatility estimated at 1.6, signifying a right heavy-tailed distribution. Again, this means that larger firms dominate the process of input resource reallocation to more efficient larger firms. In contrast, entry-exit dynamics produce the highest average productivity contribution to ALP growth at the net entry of 8.78 percent. The higher average productivity growth exhibited by the entrants' component indicates that there are extreme positive outliers pulling the mean over time

However, the standard definition of aggregate productivity growth used so far is based exclusively on establishment-level technical efficiency residuals. The literature defines this index as input-share weighted changes in the distribution of firm-level technical efficiency, decomposed into technical efficiency and technical efficiency reallocation. The latter is defined as the product between the change in the share of labour and the averaged log-level of productivity aggregated across all firms. In this thesis, we appeal to the microfoundations of aggregate productivity growth that allows for the use of labour changes weighted by the marginal product of labour to estimate labour reallocation across firms.

Therefore, the investigation began with determining primary input trends, aggregate productivity and factor-intensities in Swazi manufacturing firms over a period of trade liberalization. It proceeded with descriptive analyses and then investigated the drivers of aggregate productivity growth over time and across industries. A cross-country comparison of drivers of aggregate labour productivity growth with those of the Swazi manufacturing sector was also undertaken. We then deepened the analysis to focus on Swaziland by decomposing aggregate labour productivity growth over time using traditional methods and also relying on Petrin and Levinsohn (2012) as applied by Nishida *et al*. (2014). We

concluded with an analysis of seemingly outlying aggregate labour productivity growth in 1998 and 1999 to determine the characteristics of entrants associated with it.

The descriptive evidence showed a decline in both aggregate labour and capital productivities and an increase in the capital−labour ratio. It also showed a leftward distribution of ALP and increasing heaviness of both tails. Three potential explanations for this were identified. First, firms shed more labour relative to capital due to capital irreversibility and to South African companies shifting production back to South Africa as a response to the lifting of economic sanctions, while keeping Swazi plants in operation to cover their variable costs. Second, there was entry of lower ALP firms.

An in-depth analysis using the conventional approach found that the ALP growth was driven largely by net entry, then by cross-firm share-shift and negatively by within-firm technical change. This result was robust to controlling for confounding effects of plant turnover in the Baily *et al*. (1992) method. Using the Petrin and Levinsohn (2012) approach produced the same order of importance for APG components. That is, the net-entry contribution explained most of the changes in APG followed by input reallocation, while technical efficiency remained negative every year.

However, the most interesting case was the combined input reallocation reflecting cross-plant movements. The average reallocation of the input bundle from low to high productivity incumbent plants was 0.15 percent per year. However, isolating the average annual rate of labour reallocation from the contribution of all inputs put together produced 3.25 percent. Furthermore, paid employment showed positive growth in every year and accounted for an average of about 98 percent of all labour reallocated per year. These results are robust to 'single-deflation' by the manufacturing value-added deflator and 'double-deflation' by consumer price index. Furthermore, the annual average of net-entry contribution to APG was 58.01 percent and was mainly accounted for by the dramatic increase of APG in 1998 and 1999 due to firm entry. Thus, the analysis reveals that individual contributions by the extensive and intensive margins of resource reallocation to APG decisively *dominate* technical efficiency in the manufacturing sector in Swaziland. Firms were not investing more in improving production efficiency through innovation and adoption of new technologies than they were moving labour to higher activity producers. This conclusion remained robust regardless of the deflation procedure used in the estimation of the real value-added production function. The novelty of our results lies in the use of microfoundations to define aggregate productivity growth.

The investigation concludes with determining the presence of state dependence, unobserved heterogeneity and the impact of real sales/capital ratio on investment rate in the Swazi manufacturing. It began with descriptive analyses of a panel dataset for 13 industries and found that the rate of investment was as low as 0.24 percent every year, with the observed investment heterogeneity measured by the standard deviation just as low at 0.29.

As is typical, investment inactivity dominates the distribution of investment rates, with or without maintenance and repairs exclusions. Under conditions of new as well as maintenance and repairs investment in PME, 44.05 percent of firm-year observations experienced zero investment in the 10-year period. About 14.53 percent and 11.44 percent of these observations experienced a single and two investments in the same period, respectively. Only 36.12 percent of firm-year observations have at least two consecutive, non-missing investment sequences. There is generally a positive correlation between missingness patterns of investment and the likelihood of firms experiencing them.

The analysis of the empirical hazard function of microeconomic investment spikes confirms the lacklustre investment patterns of the manufacturing sector during the period of trade liberalization. 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. Such behavioural patterns of investment are consistent with relatively more focus on industrial M&R for machinery and equipment, and much less on new PME investment and/or replacement. This implies that the manufacturing sector in Swaziland experienced limited capital adjustment costs of installation and worker re-training. It also experienced marked obsolescence in machinery and equipment due potentially to heightened economic uncertainty.

Using a structural model of investment, we investigated the effects of past investments, unobserved heterogeneity and the 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 found 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 unobserved individual firm characteristics underlying the discrete choice to invest in PME have an insignificant effect on investment rates. This suggests that technological change does not translate into transformational and entrepreneurial investment. Consequently, 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 to 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, unless employment is controlled for. 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 withhold $t-1$ investment until period t.

Furthermore, 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.

We obtain the most interesting results when the impact of missing values of *net* PME expenditure on the investment rate are investigated in detail, for the first time as far as we are aware. 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 leads to a reduction in the probability 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*.

Since the conditional model 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 different 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 *et al.* (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 (1981c) who generalizes a transition model of binary responses that incorporates lags for both observed and latent responses. Fifth, we could also 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 (1981c), Bollen and Curran (2004) and Skrondal and Rabe-Hesketh (2014). Other possibilities include nonlinear state space models of longitudinal data, see Gala (2015), Fahrmeir and Tutz (2001, Section 4), Bayesian inference (Hasegawa, 2009) and generalized structural model equations (Skrondal and Rabe-Hesketh, 2004).

In summary, the main findings of this dissertation are that the Swazi economy experienced a deterioration of structural change and a decline in industrial aggregate productivity growth during the trade liberalization episode of the 1990s. Economic activity was characterized by a decade-long stagnation with a high probability of hollowing out in the manufacturing sector. The firm-size distribution itself in this sector evolved leftwards to a cross-sectional "missing middle" form by 2003, which is a common feature of underdevelopment in developing economies. Furthermore, there has been a complete failure of the industrial job creating prowess of small firms in Swaziland. This suggests an absence of plant-level transition channels from subsistence to transformational entrepreneurship in the Swazi manufacturing sector. At the same time, while job destruction generally dominated job creation, most of the churning involved larger firms. In the decomposition of APG, the sector revealed three sources of growth or decline. First, firm-level innovation and technical advancement; that is, technical efficiency, was APG-reducing. Second, a positive productivity growth contribution was generated by the large firm-entry margin. Third, labour reallocation from low to high productivity firms was persistently APG-enhancing. Finally, the manufacturing sector in Swaziland experienced depressed capital adjustment activities during this period of economic reforms in the Customs Union. That is, there was coexistence of both high incidence of zero and rare lumpy capital investment in PME in the manufacturing sector.

REFERENCES

Abel A.B. and O. Blanchard (1986), "The Present Value of Profits and Cyclical Movements in Investment", *Econometrica*, 54(2): pp. 249-273.

Abel A.B., A.K. Dixit, J.C. Eberly, and R.S. Pindyck (1996), "options, the Value of Capital, and Investment", *Quarterly Journal of Economics*, 111(3): pp. 573-777.

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.

Acemoglu D., V. M. Carvalho, A. Ozdaglar and A. Tahbaz-Salehi (2012), "The Network Origins of Aggregate Fluctuations", *Econometrica*, Vol. 80(5): pp. 1977–2016.

Ackerberg, D.A., K. Caves and G. Frazer (2015), "Identification Properties of Recent Production Function Estimators", *Econometrica*, Vol. 83, No. 6: pp. 2411–2451.

Aghion, P., M. Braun and J. Fedderke (2007), "Competition and productivity growth in South Africa", *Economics of Transition,* 16(4): 741-768.

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: pp. 1189-1204.

Albarran, P., R. Carrasco and J.M. Carro (2017), "Estimation of Dynamic Nonlinear Random Effects Models with Unbalanced Panels", Working Paper 15-03, Economic Series, Departamento de Economía, UC3M.

Albert, P. S. and D. A. Follmann (2003), "A random effects transition model for longitudinal binary data with informative missingness", *Statist. Neerland.*, Vol. **57**: pp. 100–111.

Alfaro L., A. Charlton and F. Kanczuck (2008), "Plant Size Distribution and Cross Country Differences", NBER WP 14060.

Alfaro, L., A. Charlton and F. Kanczuk (2008), "Firm-Size Distribution and Cross Country Income Differences." National Bureau of Economic Research Working Paper 14060.

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

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

Arulampalam, W. and M. B. Stewart (2009), "Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and Comparison with Alternative Estimators", *Oxford Bulletin of Economics and Statistics*, Vol.71 (5):pp. 659–681.

Asplund, M. and V. Nocke (2006), "Firm Turnover in Imperfectly Competitive Markets", *Review of Economic Studies* **73**, 295–327.

Axtell, R.L. (2001), "Zipf Distribution of U.S. Firm Sizes", *Science*: pp. 1818-1820.

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

Baily M, Hulten C, and Campbell D (1992), "Productivity Dynamics in Manufacturing Plants" in E. Dean, M. Harper and C. Hulten (Eds): *New Directions in Productivity Analysis*, University of Chicago Press: pp. 303-363.

Banerjee, A.V. and B. Moll (2010), Why Does Misallocation Persist? *American Economic Journal: Macroeconomics*, 2(1): pp. 189-206.

Banerjee, A.V. and E. Duflo (2005), "Growth Theory through the Lens of Development Economics." In Handbook of Economic Growth, Vol. 1 A, eds. P. Aghion and S. N. Durlauf, 473-552. Amsterdam: Elsevier.

Bartelsman E., J. Haltiwanger and S. Scarpetta (2004), "Microeconomic evidence of creative destruction in industrial and developing countries", *World Bank policy research working paper 3464*.

Bentolila, S. and G. Bertola (1990), "Firing Costs and Labour Demand: How Bad is Eurosclerosis?" *The Review of Economic Studies*, Vol.57 (3): pp381-402.

Bernard A.B., J.B. Jensen, S.J. Redding and P.K. Scott (2016), "Global Firms", NBER WP 22727.

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).

Birch, D. (1987), "Job Creation in America: How Our Smallest Companies Put the Most People to Work". New York: Free Press.

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

Bloom N., B. Eifert, A. Mahajan, D. McKenzie and J. Roberts (2013), "Does Management Matter? Evidence from India", *The Quarterly Journal of Economics*, V. 128(1): pp. 1-51.

Bloom N., R. Lemos, R. Sadun, D. Scur and J. Van Reenen (2014), "The New Empirical Economics of Management", *NBER Working Paper Series*.

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

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

Bollen, K.A. and P.J. Curran (2004), "Autoregressive latent trajectory (ALT) models: a synthesis of two traditions", *Sociological Methods & Research*, Vol.32: pp. 336–383.

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

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

Bowsher, C.G. (2002), "On Testing Overidentifying Restrictions in Dynamic Panel Data Models", Mimeograph.

Brixiová, Z. and T. Kangoye (2013), "Youth Employment in Africa: New Evidence and Policies from Swaziland", *Discussion Paper No. 7467*.

Brown, J.D. and J. Earle (2002), "Understanding the Contributions of Reallocation to Productivity Growth: Lessons from a Comparative Firm-Level Analysis", *Discussion Paper 2008/05*.

Browning M. and J, Carro (2010), "Heterogeneity in Dynamic Discrete Choice Models", *The Econometrics Journal*, 13(1): pp. 1-39.

Bruno, M. (1978), "Duality, Intermediate Inputs and Value-Added" in Fuss M. and D. McFaden, Eds: Production Economics: A Dual Approach to Theory and Applications, Vol.2, Amsterdam.

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.

Butani S.J., R.L. Clayton, V. Kapani, J.R. Spletzer, D.M. Talani and G.S. Werking (2006), "Business Employment Dynamics: Tabulations by Employer Size", Monthly Labour Review", 129(2): pp. 3-22.

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. and M.L. Hammour (1994), "The Cleansing Effect of Recessions", *The American Economic Review*, Vol. 84, No. 5, pp. 1350-1368.

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, Cambridge University Press.

Cameron, A.C. and P.K. Trivedi (2009), "Microeconometrics using Stata", College Station, TX: Stata Press.

Central Bank of Swaziland (1995−2003), Annual Reports.

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", Mimeograph.

Christen, P. (2012), "Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection", Berlin, Heidelberg, New York: Springer-Verlag.

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: pp. 1-45.

Chrysanthou, G.M. (2008), "Determinants of Trade Union Membership in Great Britain During 1991-2003", Mimeograph.

Cooper, R.W. and J.C. Haltiwanger (2006), "On the Nature of Capital Adjustment Costs", The Review of Economic Studies, Vol. 73(3): pp. 611-633.

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): pp. 921-946.

Dasgupta K. (2015), "The Missing Middle in Developing Countries Revisited", Mimeograph, University of Toronto.

Dasgupta, K. (2010), "The Missing Middle in Developing Countries Revisited", Mimeograph.

Davidsson P., L. Lindmark and C. Olofsson (1998), "The extent of overestimation of small firm job creation- an empirical examination of the "regression bias"", *Small Business Economics Vol.*10: pp. 87-100.

Davis S. and J. Haltiwanger (1992), "Gross Job Creation, Gross Job Destruction and Employment Reallocation", *Quarterly Journal of Economics*, 107(3): pp. 819-863.

Davis S., and J. Haltiwanger (1996), "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts", *Small Business Economics*, 8(4): pp. 297-315.

Davis, S. J., J. Haltiwanger, and S. Schuh (1996a), Job Creation and Destruction. Cambridge, MA: MIT Press.

Davis, S.J. and J. Haltiwanger (1999), "Gross Job Flows" in O. Ashenfelter and D.Card Eds: *Handbook of Labour Economics*, Vol.3: pp. 2711-2805.

De Loecker, J. and J. Konings (2006), "Job reallocation and productivity growth in a post-socialist economy: Evidence from Slovenian manufacturing", *European Journal of Political Economy*, Vol. 22: pp. 388–408.

Decker R.A, J. Haltiwanger, R.S. Jarmin, J. Miranda (2016), "Where has all the skewness gone? The decline in high-growth (young) firms in the U.S.", *European Economic Review*, 86: pp.4–23.

Doms M., T. Dunne and K.R. Troske (1997), "Workers, Wages and Technology", *Quarterly Journal of Economics*, 112(1): pp. 253-290.

Doraszelski, U. and J. Jaumandreu (2013), "R&D and productivity: Estimating endogenous productivity", *Review of Economic Studies*, Vol.80: pp. 1338–1383.

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.

Dunne, T., M.J. Roberts and L. Samuelson (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries", *The RAND Journal of Economics*, Vol. 19(4): pp. 495-515.

Dutoit, L.C. (2007), "Heckman's Selection Model, Endogenous and Exogenous Switching Models: A survey", Mimeograph.

Eberly, J., S. Rebelo and N. Vincent (2012), "What Explains the Lagged Investment Effect?" Mimeograph.

Edwards L., F. Flatters, M. Stern and Y. Ramskolowan (2013), "Swaziland Economic Diversification Study: Final Report", *DNA Economics*.

Edwards, L and A. Behar (2005), "Trade liberalisation and labour demand within South African manufacturing firms", Working Paper Number 6.

Edwards, L. (2006), "Has South Africa Liberalised its Trade?", Trade and Poverty Project, Southern African Labour and Development Research Unit, University of Cape Town.

Erickson, T. and T.M. Whited (2000), "Measurement Error and the Relationship between Investment and *q*", *Journal of Political Economics*, 108(5): pp. 1027-1057.

Ericson, R. and A. Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *The Review of Economic Studies*, Vol. 62, No. 1, pp. 53-82.

Escribano, A. and R. Stucchi (2014), "Does recession drive convergence in firms' productivity? Evidence from Spanish manufacturing firms", *Journal of Productivity Analysis*, 41: pp. 339–349

Faggio G. and J. Konings (2003), "Job Creation, Job Destruction and Employment Growth in Transition Economies in the 90s", *Economic Systems*, 27: pp. 129-154.

Fahrmeir, L. and G. Tutz (2001), *Multivariate Statistical Modelling based on Generalized Linear Models*, 2nd edition, New York: Springer.

Fan J. (1992), "Design Adaptive Nonparametric regression", *Journal of the American Statistical Association*, 87(420): pp. 998-1004.

Fedderke, J., C. Kularatne and M. Mariotti (2005), "Markup Pricing in South African Industry", Department of Economics, University of California.

Foster, L., J. Haltiwanger and C. Krizan (2001), Aggregate productivity growth: Lessons from microeconomic evidence, NBER chapters, pp. 303–372.

Foster, L., J. Haltiwanger and C. Syverson (2008), "Reallocation, firm turnover, and efficiency: selection on productivity or profitability?", *The American Economic Review*, 98(1): pp. 394–425.

Francis, B.J., D.N. Stott and R.B. Davies (1996), *SABRE: a Guide for Users, Version 3.1*. Lancaster: Lancaster University.

Friedman, M. (1992), ''Do Old Fallacies Ever Die?'' *Journal of Economic Literature*, Vol.30 (1992): pp. 2129–2132.

Gabaix, X. (2011), "The Granular Origins of Aggregate Fluctuations", *Econometrica*, Vol. 79 (3): pp. 733–772.

Gala, V. (2015), "Measuring Marginal q", Mimeograph.

Galuščák, K. and L. Lizal (2011); "The Impact of Capital Measurement Error Correction on Firm-Level Production Function Estimation", CNB Working Paper Series.

Gandhi A., S. Navarro and D. Rivers (2016), "On the Identification of Production Functions: How Heterogeneous is Productivity?", Mimeograph.

Gebreeyesus, M. (2008); "Firm turnover and productivity differentials in Ethiopian manufacturing", *Journal Productivity Analysis,* Vol. 29, No. 2, Special issue on Transitioning Economics : pp. 113-129.

Gelb, A., C.J. Meyer and V. Ramachandran (2014), "Development as Diffusion: Manufacturing Productivity and Sub-Saharan Africa's Missing Middle", *Centre for Global Development*, Working Paper 357.

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, S and C. 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.

Griliches, Z. and H. Regev (1995), "Firm productivity in Israeli industry", 1979-1988, *Journal of Economics*, 65(1):pp. 175–203.

Hajivassiliou, V.A. and Y.M. Ioannides (2007), Unemployment and liquidity constraints. *Journal of Applied of Econometrics*, Vol.22: pp**.** 479–510.

Haltiwanger, J. (1997), "Measuring and Analysing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence", *Review*: 55-77.

Haltiwanger, J., R.S. Jarmin and J. Miranda (2013), "Who Creates Jobs? Small Versus Large Versus Young", *The Review of Economics and Statistics*, VOL. XCV(2).

Hamermesh, D.S., W. H. J. Hassink and J. C. van Ours (1996), "Job Turnover and Labour Turnover: A Taxonomy of Employment Dynamics", *The Microeconometrics of Human-Resource Management: Multinational Studies of Firm Practices*, pp. 21-40.

Hammouda, H.B., S.N. Karingi, A.E. Njuguna and M.S. Jallab (2010); "Growth, productivity and diversification in Africa", *Journal Productivity Analysis*, 33:125–146.

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.

Hartigan J.A., and P.M. Hartigan (1985), "The Dip Test of Unimodality", *The Annals of Statistics*, 13(1): pp. 70-84.

Hasegawa, H. (2009), "Bayesian dynamic panel-ordered probit model and its application to subjective well-being", *Communications in Statistics - Simulation and Computation*, Vol. 38: pp. 1321–1347.

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", Mimeograph, Hiroshima University.

Hayakawa, K. and M.H. Pesaran (2015), "Robust standard errors in transformed likelihood estimation of dynamic panel data models with cross-sectional heteroskedasticity", *Journal of Econometrics*, Vol. 188: pp. 111–134.

Hayashi F. (1982), "Tobin's Marginal q and Average q: A Neoclassical Interpretation", *Econometrica*, 50(1): pp. 213-224.

Hayashi F. (2000), "Econometrics", Princeton University Press, New Jersey.

Heckman J.J. and B. 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. (1981c), "Statistical models for discrete panel data", 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. (1981c), Statistical models for discrete panel data. 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. and B. Singer (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52 (2): pp. 271-320.

Hijzen, A., R. Upward and P.W. Wright (2010), "Job Creation, Job Destruction and the Role of Small Firms: Firm-Level Evidence for the UK", *Oxford Bulletin of Economics and Statistics*, 72(5): pp. 621-647.

Ho, A. T. Y., K.P. Huynh, D.T. Jacho-Chávez and G. Cubas (2014), "Productivity, Reallocation, and Economic Crisis: Evidence from Ecuadorian Firm-Level Data", Mimeograph.

Honoré B.E. and E. Kyriazydou (2000), "Panel Data Discrete Choice Models with Lagged Dependent Variables", *Econometrica*, 68(4): pp. 839-874.

Horowitz, J. (2001), "The Bootstrap": in J. J. Heckman and E. Leamer (eds.) *Handbook of Econometrics,* Vol. 5, (Oxford: Elsevier Science): pp. 3159-3228.

Hsiao, C. (2003), "Analysis of Panel Data", Econometric Society Monographs, Cambridge University Press, 2nd ed.

Hsieh, C.T. and B.A. Olken (2014), "The Missing "Missing Middle"", *Journal of Economic Perspectives*, Vol.28 (3): pp.89-108.

Hsieh, C.T. and P.J. Klenow (2009), "Misallocation and Manufacturing TFP in China and India", *The Quarterly Journal of Economics*, V.124(4): pp. 1403-1448.

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): pp. 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.

Hurst, E. and B.W. Pugsley (2011), "What Do Small Businesses Do?", Mimeograph.

Hyslop, D.R. (1999), "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labour Force Participation Behaviour of Married Women," *Econometrica*, 67: pp. 1255-1294.

IMF Country Report (2000).

Isaksson, A. (2010), "Structural Change and Productivity Growth: A Review with Implications for Developing Countries", Research and Statistics Branch, Programme Coordination and Field Operations Division, UNIDO.

Jarmin, R.S. and J. Miranda (2002), ''The Longitudinal Business Database'' CES working paper 02:17 (2002).

Jeong, H. and R. M. Townsend (2007), "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, Vol.32 (1): pp.179-221.

Jonsson, G. and A. Sumbranian (2001), "Dynamic Gains from Trade: Evidence from South Africa", *IMF Papers*.

Jovanovic, B. (1982), "Selection and Evolution of Industry", *Econometrica*, Vol.15 (3): pp. 649-670.

Judson, R.A. and A.L. Owen (1999), "Estimating dynamic panel data models: a guide for macroeconomists", *Economics Letters*, 65: pp. 9–15.

Kalbfleisch J.D. and R. L. Prentice (2002), *The Statistical Analysis Failure Time Data*, John Wiley & Sons, Hoboken, New Jersey.

Kambourov, G. (2009), "Labour Market Regulations and the Sectoral Reallocation of Workers: The Case of Trade Reforms", *Review of Economic Studies,* Vol. 76: pp. 1321-1358.

Kapetanios, G. and M. Marcellino (2007), "Factor-GMM Estimation with Large Sets of Possibly Weak Instruments", Mimeograph.

Kerr, A., M. Wittenberg and J. Arrow (2013), "Job Creation and Destruction in South Africa", A Southern Africa Labour and Development Research Unit Working Paper Number 92. Cape Town: SALDRU, University of Cape Town.

Khan A. and J.K. Thomas (2008), "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics", *Econometrica*, 76(2): pp. 395-436.

King and Nielson (2016), "Why Propensity Scores Should Not Be Used for Matching", Mimeograph.

Klapper, L. and C. Richmond (2011), "Patterns of business creation, survival and growth: Evidence from Africa," *Labour Economics*, Vol.18 (S1): S32-S44.

Konings, J., H. Lechmann and M.E. Schaffer (1996), "Job Creation and Job Destruction in a Transfer Economy: Ownership, Firm Size, and Gross Job Flow in Polish Manufacturing 1988-91", *Labour Economics*, Vol.3: pp. 299-317.

Krueger, N.O. (2013), "The Missing Middle" Chap. 9 in *Economic Reform in India: Challenges, Prospects, and Lessons* edited by Nicholas C. Hope, Anjini Kochar, Roger Noll, and T.N. Srinivasan, Cambridge University Press.

Kwon, H., F. Narita and M. Narita (2009), "New Evidence on Japan's Aggregate Productivity Growth in the 1990s", Mimeograph.

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 G.A. Pfann (2007), "Structural Identification of High and Low Investment Regimes", *Journal of Monetary Economics*, 54: pp. 797–819.

Levinsohn, J. and A. Petrin (1999), "When Industries Become More Productive, Do Firms?: Investigating Productivity Dynamics", Mimeograph.

Levinsohn, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies,* 7: 317-341.

Levinson, M. (2016), "U.S. Manufacturing In International Perspective", Congressional Research Service, CRS Report Prepared For Members And Committees Of Congress.

Lewis, W.A. (1954), "Economic Development with Unlimited Supplies of Labour," *Manchester School of Economic and Social Studies*, vol. 22: pp. 139-191.

Lileeva, A. (2008), "Trade Liberalization and Productivity Dynamics: Evidence from Canada", *The Canadian Journal of Economics*, Vol.41(2): pp. 360-390.

Lindsay, B.G., (1983), "The Geometry of Mixture Likelihoods: A General Theory," *The Annals of Statistics*, 11: pp. 86-94.

Liu, L. and J.R. Tybout (1996), "Productivity Growth in Chile and Colombia: The Role of Entry, Exit and Learning", in M.J. Roberts and J.R. Tybout, Eds.: Industrial Revolution in Developing Countries, The World Bank, Oxford University Press.

Lokshin, M. and Z. Sajaia (2004), "Maximum likelihood estimation of endogenous switching regression models", *The Stata Journal*, 4(3): pp. 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 Dependent and Qualitative Variables in Econometrics", New York: Cambridge University Press.

Masuku, M.B. and T.S. Dlamini (2009), "Determinants of Foreign Direct Investment in Swaziland", *Journal of Development and Agricultural economics*, Vol. 1(5): pp. 177-184.

Mazumdar, D. and S. Sarkar (2008), "The Employment Problem in India and the Phenomenon of the

McClelland J. (1997), "Permit Renewal and the Timing of Major Investment", Mimeograph, University of Maryland.

McMillan, M. and D. Rodrik (2014), "Globalization, Structural Change And Productivity Growth", *NBER Working Paper Series*, Working Paper 17143.

McMillan, M. and K. Harttgen (2014), "What is Driving the 'African Growth Miracle'?" *African Development Bank Group*, Working Paper No. 209.

McMillan, M., D. Rodrik and Verduzco-Gallo (2014), "Globalization, Structural Change, and Productivity Growth, with an Update on Africa", *World Development*, Vol. 63: pp. 11–32.

Mehrhoff, J. (2009), "A solution to the problem of too many instruments in dynamic panel data GMM", *Discussion paper n. 1/2009*, Deutsche Bundesbank.

Melitz, M. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, Vol. 71(6): pp. 1695-1725.

Memedovic O. and L. Apadre (2010), "Structural Change in the World Economy: Main Features and Trends", Research and Statistics Branch, Working Paper 24/200.

Ministry of Economic Planning and Development Report (2004/2005), Economic Development Plan.

Mulally, Alan R. (2008), "Review of industry plans to stabilize the financial condition of the American automobile industry." Hearing, *United States House of Representatives Committee on Financial Services*.

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.

Nassif, A., C. Feijó and E. Araújo (2014), "Structural change and economic development: is Brazil catching up or falling behind?", *Cambridge Journal of Economics*: pp. 1 - 26.

Neumark, D., B. Wall and J. Zhang (2011), "Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series", *The Review of Economics and Statistics*, 93: pp. 16-29.

Nielson, Ø. A. and F. Schiantarelli (2003), "Zeros and Lumps in Investment: Empirical Evidence on Irreversibilities and Nonconvexities", *The Review of Economics and Statistics*, Vol. 85, No. 4, pp. 1021-1037.

Nishida, M., A. Petrin and S. Polanec (2014), "Exploring reallocation's apparent weak contribution to growth", *Journal Productivity Analysis,* Vol. 42(2): pp.187-210

OECD (2001), "Measurement of Aggregate and Industry-Level Productivity Growth", *OECD Manual*.

Olley, S., and A. Pakes (1996), "The dynamics of productivity in the telecommunications equipment industry", *Econometrica,* 64: pp. 1263– 1298.

Pagés C., G. Pierre and S. Scarpetta (2008), "Job Creation in Latin America and the Caribbean: Recent Trends and Policy Challenges", *The International Bank for Reconstruction and Development / The World Bank*.

Pakes, A. and R. Ericson (1998), "Empirical Implications of Alternative Models of Firm Dynamics", *Journal of Economic Theory*, 79: pp. 1-45.

Pavcnik, N. (2002), "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants", *The Review of Economic Studies*, Vol. 69(1): pp. 245-276.

Petrin, A. and J. Levinsohn (2012), "Measuring aggregate productivity growth using plant-level data"**,** *RAND Journal of Economics*, Vol. 43(4): pp. 705–725.

Petrin, A. and J. Sivadasan (2013), "Estimating Lost Output from Allocative Inefficiency, with an Application to Chile and Firing Costs", *The Review Of Economics And Statistics,* 95(1): pp. 286–301.

Petrin, A., B.P. Poi and J. Levinsohn (2004), "Production function estimation in Stata using inputs to control for unobservables", *The Stata Journal*, 4(2): pp. 113–123.

Petrin, A., T.K. White and J.P. Reiter (2011), "The impact of plant-level resource reallocations and technical progress on U.S. macroeconomic growth", *Review of Economic Dynamics*, 14: pp. 3–26

Pham H.T. and S. Takayama (2015), "Revisiting the "Missing Middle": Productivity Analysis", Mimeograph.

Pudney, S. (2008), "The dynamics of perception: modelling subjective wellbeing in a short panel" *Journal of the Royal Statistical Society,* Vol. A.171: pp. 21–40.

Quah, D.T. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Centre for Economic Discussion*, Discussion Paper No.324.

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.

Ranis, G. and J.C. Fei (1961), "A Theory of Economic Development," *American Economic Review*, vol. 51, September 1961, pp. 533-558.

Restuccia, D. and R. Rogerson (2008), "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics*, 11(4): pp. 707-20.

Restuccia, D. and R. Rogerson (2008), "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics*, 11(4): pp. 707-20.

Restuccia, D. and R. Rogerson (2012), "Misallocation and Productivity", Mimeograph.

Roberts, M.J. and J.R. Tybout (1996), "Industrial Evolution in developing Countries: Micro Patterns of Turnover, Productivity and Market Structure", World Bank.

Rodrik, D. (2006a), "Industrial Development: Stylized Facts and Policy Directions" in *Industrial Development for the 21st Century: Sustainable Development Perspectives,* United Nations, New York.

Rodrik, D. (2006b), "Understanding South Africa's Economic Puzzles", Harvard University Center for International Development Project on South Africa.

Rodrik, D. (2013), "Structural Change, Fundamentals, and Growth: An Overview", Mimeograph.

Rodrik, D. (2014), "An African Growth Miracle?", NBER Working Paper Series, Working Paper 20188.

Rodrik, D. (2015), "Premature Deindustrialization", *NBER Working Paper Series*, Working Paper 20935.

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.

Rossi-Hansberg E. and M.L.J. Wright (2007), "Establishment Size Dynamics in the Aggregate Economy", *The American Economic Review*, 97(5): pp. 1639-1666.

Sandefur, J. (2010), "On the Evolution of the Firm Size Distribution in an African Economy", CSAE WPS/2010-05.

Sandefur, J. and A. Glassman (2013), "The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics", Mimeograph.

Schoar, A. (2010), "The Divide between Subsistence and Transformational Entrepreneurship", Mimeograph.

Seaman S., J. Galati, D. Jackson and J. Carlin (2013), "What is meant by "Missing at Random"?" *Statistical Science*, 28(2): pp. 257–268.

Shea J. (1997), "Instrument Relevance in Multivariate Linear Models: A Simple Measure", *Review of Economics and Statistics*, 79(2): pp. 348-352.

Shiferaw, A. and A. Bedi (2009), "The Dynamics of Job Creation and Job Destruction: Is Sub-Saharan Africa Different?", IZA DP No. 4623.

Skrondal, A. and S. Rabe-Hesketh (2004), *Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton: Chapman and Hall–CRC.

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.

SME report (2003), "A Survey of Small and Medium Enterprises in Swaziland", Government of Swaziland.

Solow, R.M. (1957), "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics,* Vol. 39: pp. 312–320.

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.

Stock J.H. and J.H. Wright (2000), "GMM with Weak Identification", *Econometrica*, 68(5): pp. 1055-1096.

Stock J.H. and M. Yogo (2005), "Testing for Weak Instruments in Linear IV Regression" in D.W.K. Andrews and J.H. Stock (eds): *Identification and Inference in Econometric Models*, Cambridge: pp. 80-108.

Syverson, C. (2004), "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, 86(2): pp. 534–50.

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.

Teal F. (2016), "Firm Size, Employment and Value Added in African Manufacturing Firms: Why Ghana Needs its 1 Percent", Centre for the Study of African Economies, WPS/2016-2017, University of Oxford.

Tibiyo TakaNgwane (2010, 2012), Annual Reports.

Timmer, M., G. de Vries and K. de Vries (2014), "Patterns of Structural Change in Developing Countries", Mimeograph.

Tybout J.R. (2000), "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?", *Journal of Economic Literature, Vol. XXXVIII: pp.11–44.*

Tybout J.R., (2014), "The Missing Middle, Revisited", Pennsylvania State University and NBER, Mimeograph.

Van Biesebroeck (2007), "Robustness of Productivity Estimates", *The Journal of Industrial Economics*, Vol. 55 (3): pp. 529-569.

Van Biesebroeck, J. (2005), "Firm Size Matters: Growth and Productivity Growth in African Manufacturing", *Economic Development and Cultural Change*, 53(3): pp. 545-583.

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. (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. (2001), "Applications of Generalized Method of Moments Estimation", *The Journal of Economic Perspectives*, Vol. 15(4): pp. 87-100.

Wooldridge, J.M. (2009), "On estimating firm-level production functions using proxy variables to control for unobservables", *Economics Letters*, 104: pp. 112–114.

Wooldridge, J.M. (2010), "Correlated Random Effects Models with Unbalanced Panels", Mimeograph.

Young, A. (2012), "The African Growth Miracle", *Journal of Political Economy*, Vol. 120(4): pp. 696-739.

Youssef, A.H. and M.R. Abonazel (2015), "Alternative GMM Estimators for First-Order Autoregressive Panel Models: An Improving Efficiency Approach", MPRA.

Youssef, A.H., A.A. El-Sheik and M.R. Abonazel (2014), "Improving the Efficiency of GMM Estimators for Dynamic Panel Models", MPRA.

Ziliak, J.P. (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.