Chapter 5

Linkages at Work? Empirical Results

The history of South African industrial development and the results of the simulation both provided evidence that linkages between different sectors might influence industrial development in the country. However, neither provide 'hard' empirical evidence about their actual strength and impact. In order to rectify this omission, I propose to use a structural vector auto-regression model (SVAR) to assess sectoral development in South Africa from 1970 to 2006. The model explicitly tests for the impact of input-output linkages on the growth performance of sectors.

The chapter starts with an overview of existing empirical work in the New Economic Geography field and on linkage effects, particularly those studies with direct relevance for South Africa. In the second part, the SVAR methodology is introduced in detail. The third part presents the model and estimation results. I conclude with an interpretation of the results.

5.1 Empirics of Linkage Effects

Existing empirical applications of NEG models are reviewed briefly below. Since there are a number of problems in these applications and since this study has a broader focus on linkage effects that does not limit itself to the model results, I also add a brief description of studies that attempt to assess linkage effects in South Africa.
5.1.1 Empirics of the New Economic Geography

Turning to empirical applications of NEG first, it is noteworthy that they still lag their theoretical counterparts in terms of quantity and coherence, despite a recent surge in publications (Head and Mayer, 2004, 2611 and Brakman and Garretsen, 2006, 569). In his critical assessment of NEG, Martin finds this to be embedded in the very nature of the models that to him “[...] do not lend themselves easily to empirical estimation or application, since they are typically too abstract, over simplified and too idealised.” (Martin, 1999, 70) Empirical work therefore mostly limits itself to the testing of certain properties of the models.

Krugman offered first insights into agglomeration effects by calculating locational Gini coefficients as an index of industry concentration (Krugman, 1991b, 54ff.). These coefficients compare the incidence of particular industries in regions to total manufacturing production. Krugman finds high concentration for most US industries. Ellison and Glaeser substantially enrich this concept in two contributions (Ellison and Glaeser, 1997, 1999). They measure industry concentration against an accidental distribution and still find concentration in most sectors. Yet they caution that often, agglomeration will result from natural advantages (Ellison and Glaeser, 1997). In their second paper they try to empirically isolate the effect of natural advantages (or first nature) on concentration and find that “about 20% of observed geographic concentration can be explained by a small set of advantages.” (Ellison and Glaeser, 1999, 315). Apart from indices of concentration, Head and Mayer (2004) provide an exhaustive overview of empirical tests on various propositions of the New Economic Geography.

Regarding linkage effects in particular, Ellison and Glaeser (1997) measure coagglomeration of industries with strong upstream-downstream relationships. For the constructed pairs of industries 77 out of 100 downstream industries that are heavily using one input are locally coagglomerated with the upstream producer (Ellison and Glaeser, 1997, 917) which is a strong indication that vertical linkage effects are indeed at play in spatial agglomeration processes. Building on the multi-sector VL model presented in Puga and Venables (1996) and in Fujita et al. (2001), Forslid et al. (2002) calibrate a 14 sector model with a complete input-output structure derived from IO-matrices to study the effects of economic integration, which they simulate by a fall in trade costs over time, on industry concentration for four European regions.
However, these results have to be interpreted with care. Hypotheses derived from NEG models are often equivalent to hypotheses derived from other theories—thus their empirical testing and corroboration is “not necessarily evidence of the relevance of NEG, but could support competing theories and hypotheses.” (Brakman and Garretsen, 2007, 99) Ottaviano (2007) points to a second identification problem: the ‘within-equivalence’ of vertical linkages models and those NEG models based on factor mobility (e.g. the core-periphery model). He shows that the ‘footloose entrepreneur’ (FE) model, a variation of the core-periphery model, has the same structural features and equivalent equilibrium properties to the vertical linkages footloose entrepreneur (VLFE) model, a variation of the vertical linkages models. Thus, observed agglomeration could stem from either labour mobility as predicted in the core-periphery setting or from linkages, but it is not possible to uniquely attribute it to one of the two effects (Ottaviano, 2007, 64).

To my knowledge, there is no empirical work investigating agglomeration due to linkage effects in South Africa. However, a number of contributions try to estimate concentration tendencies with respect to manufacturing industries in South Africa and the Southern African region, in particular the Southern African Development Community (SADC). McCarthy provides a qualitative assessment of the potential effects of a free trade area established in SADC and suggests that some polarisation in favour of South Africa due to agglomeration effects as described in the new economic geography is to be expected (McCarthy, 1999). For a similar purpose, Petersson (2002) calculates centrality and concentration indices for SADC countries and South African provinces. He finds strong evidence for a core periphery structure (with South Africa as the regional core and the Gauteng region as core region of South Africa) and concentration of production in industry, particularly in labour and scale intensive sectors. In more recent work, Hess (2005) uses the locational Gini coefficient to establish that while some industries tend to concentrate in the core of SADC, there is also some concentration in peripheral countries (e.g. textiles and wearing apparel). Heavy industries such as industrial chemicals, iron and steel and non-ferrous metals are increasingly located in South Africa however (Hess, 2005, 52). Fedderke and Szalontai (2005) re-establish these results for manufacturing sectors within South Africa and also find that the high levels of concentration are harmful for output and productivity growth.
5.1.2 Linkage Effects in South Africa

While there is a shortage of empirical work on South Africa with explicit reference to the NEG framework, that cannot be said about linkage effects in general.

Fine and Rustomjee (1996) base their seminal work on industrial development on detailed examinations of input-output linkages, particularly with reference to the MEC. They show that mining and closely related manufacturing industries are a cohesive unit, with almost 60% of MEC-inputs stemming from the MEC itself and close to 30% of its output returned to the MEC (Fine and Rustomjee, 1996, 81). Non-MEC manufacturing on the other hand displays only weak linkages to the MEC.

The importance of linkages has also been stressed in a number of sectoral case studies recently. Roberts (2006, 166) analyses the metal products sector and finds that the most important input—basic iron and steel from within the confines of the MEC—is largely supplied by Iscor, the South African steel monopolist. Thus, the competitiveness of firms in the sector depends to a large extent on the pricing of this key input. Studies of the chemicals and plastics sector (see Roberts, 2001; Dobreva, 2006) reach similar conclusions: the downstream sector, in this case plastic firms, are strongly using chemicals as an input, and again suffer from the market power of large upstream producers.

The most recent and also most exhaustive work in this vain, which is not limited to sector case studies but aims to provide an overview over the whole economy is Tregenna (2007). Her study shows total backward linkages are strongest in the manufacturing sector (using a standard-type classification), thus the provision of an additional unit of output of manufacturing goods would require the production of 2.14 additional units of total output. Backward linkage effects for agriculture, mining and services are significantly lower (1.89, 1.76, and 1.81 respectively) (Tregenna, 2007, 105). Tregenna also calculates employment multipliers (how many jobs would be created if final demand for a sector’s output increases by one million Rand) and finds them to be higher in services than in manufacturing. These results have to be interpreted carefully however due to the well-known problems in the measurement of employment in South Africa. Unfortunately she does not report results for more deeply disaggregated data, which would be interesting from the perspective of this study. She also uses the standard distinction.
between manufacturing and mining activities, and thus does not distinguish between MEC and non-MEC manufacturing.

5.2 Testing for Linkage Effects—A SVAR Approach

Rather than testing for agglomeration effects, and thus inviting the identification problems referred to above, I choose a different, indirect strategy to analyse the significance of linkage effects. Following a paper by Abeysinghe and Forbes (2005), who engage in a similar exercise with regards to international trade relations, a structural VAR is developed that tests the effects of a shock in one sector on the other sectors in the economy. In the transmission of shocks, the role of linkages is explicitly catered for.

I will first introduce the vector auto-regression (VAR) and the structural VAR method, and then describe the specific model, data and results in detail.

5.2.1 Structural VARs

Originally, econometricians treated time series data not any different from other data and formulated traditional regression models. However, non-stationarity of many economic variables led to 'spurious' results and triggered the development of a different approach based on time series analysis—the Box-Jenkins approach (Kennedy, 2003, 319f.). Instead of relying on explanatory variables deducted from economic theory, the past values of the variable in question were used to predict its future development. “Thus in essence it is a sophisticated method of extrapolation.” (ibid.).

The dependent variable is first differentiated sufficiently to render it stationary. Then it is expressed in its own past values and current and past error terms—leading to the acronym ARIMA (autoregressive integrated moving average). Once the appropriate model (lag length, number of differentiations) is selected, it can be used to forecast the dependent variable. Because ARIMA-models outperformed traditional economists’ formulations, they were adapted and extended to a multivariate setting, producing the so-called VAR models.
Vector Autoregression

The vector autoregression approach was introduced by Sims (1980). He stated that existing large simultaneous equation models “contain too many incredible restrictions” (ibid., 14) and therefore introduced the VAR approach, where all variables in the system are endogenous and depend on their own and all the other variables’ lagged values. There are thus no restrictions introduced a priori, which also implies that the approach is atheoretical.

The variables can then be written down as a vector which is explained by its own lagged values plus an error vector. The characteristics and limitations of the VAR approach are now explained in more detail by using the example of a bivariate system with one lag. This presentation follows the outline in Enders (1995, 294ff.).

Assume that the time paths of the two variables $y_t$ and $z_t$ mutually affect each other. Then

$$y_t = b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \epsilon_{yt}$$  \hspace{1cm} (5.1)

$$z_t = b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \epsilon_{zt}$$  \hspace{1cm} (5.2)

Assume further that $y_t$ and $z_t$ are both stationary, and that the error terms $\epsilon_{yt}$ and $\epsilon_{zt}$ are uncorrelated white noise disturbances with standard deviations $\sigma_y$ and $\sigma_z$. In the above equations, there is a contemporaneous effect of $y_t$ on $z_t$ and vice versa—their size being expressed in $b_{12}$ and $b_{21}$, respectively. The epsilons are pure shocks in the two variables. However, these equations are not reduced-form equations. The endogenous variables of the system ($y_t$ and $z_t$) appear on the right hand side as well, thus violating a core assumption of linear regressions. The reason for this violation is that these regressors are correlated with the error terms due to the interdependence of the various endogenous variables that are determined simultaneously (Kennedy, 2003, 180). To lend themselves to estimations, they have to be rewritten in their reduced form, so that every endogenous variable is expressed as a function of exogenous (including lagged endogenous) variables.

Using matrix algebra, 5.1 and 5.2 become

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{yt} \\ \epsilon_{zt} \end{bmatrix}$$  \hspace{1cm} (5.3)
or in short form: \( Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \epsilon_t \). This can be premultiplied by \( B^{-1} \) to get \( x_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1}\epsilon_t \). Simplifying notation allows us to rewrite this equation as \( x_t = A_0 + A_1 x_{t-1} + \epsilon_t \) or, expressed in explicit form,

\[
y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + \epsilon_{1t} \tag{5.4}
\]

\[
z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + \epsilon_{2t} \tag{5.5}
\]

where \( a_{ij} \) are the elements of \( A_0, a_{ij} \) of matrix \( A_1 \), and \( \epsilon_{it} \) of vector \( \epsilon_t \). This is the so-called standard or reduced form of the VAR, and equations 5.4 and 5.5 can be estimated. Note however that the error terms \( \epsilon_{it} \) are now composed of the two original shocks \( \epsilon_{yt} \) and \( \epsilon_{zt} \). Thus, without further restrictions, the impact of one of these pure shocks on the dependent variables cannot be identified. Formally,

\[
B^{-1} = \frac{1}{1 - b_{12} b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}
\]

Thus, \( B^{-1}\epsilon_t \) reveals the error terms

\[
\epsilon_{1t} = \frac{\epsilon_{yt} - \epsilon_{zt} b_{12}}{1 - b_{12} b_{21}} \tag{5.6}
\]

\[
\epsilon_{2t} = \frac{\epsilon_{zt} - \epsilon_{yt} b_{21}}{1 - b_{12} b_{21}} \tag{5.7}
\]

As before, \( b_{12} \) and \( b_{21} \) are the parameters that indicate the contemporaneous effect of one variable on the other. If they are non-zero, then the structural shocks cannot be identified from the estimated (standard form) VAR. The errors \( \epsilon_{it} \) do have the desirable properties of being stationary, i.e. having zero means, a constant variance and zero autocovariances (Enders, 1995, 296) and thus do fulfill the preconditions for an OLS estimation. On the other hand, their covariance will typically be non-zero, so the two error terms will be correlated.

**Identification**

5.1 and 5.2 cannot be estimated directly, but 5.4 and 5.5 can be. OLS will yield estimates for the parameters \( a_{i0} \) and \( a_{ij} \) and we can calculate variances and the covariance of the residuals, \( \text{var}(\epsilon_{it}) \) and \( \text{cov}(\epsilon_{1t}, \epsilon_{2t}) \). But the nine estimates are not enough to identify the original equations that have ten
unknowns, \( b_{i0}, \gamma_{ij}, \) and the standard deviations \( \sigma_y \) and \( \sigma_z \). Identification requires additional restrictions of the original system.

One possible restriction is to set one of the coefficients of 5.1 and 5.2 equal to zero. If, for example, \( b_{21} = 0 \), then \( y_t \) has no contemporaneous effect on \( z_t \), and the system is exactly identified. The original time series can be rewritten as

\[
y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \epsilon_{yt} \tag{5.8}
\]
\[
z_t = b_{20} + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \epsilon_{zt} \tag{5.9}
\]

The error terms of the reduced form become

\[
e_{1t} = \epsilon_{yt} - \epsilon_{zt}b_{12} \tag{5.10}
\]
\[
e_{2t} = \epsilon_{zt} \tag{5.11}
\]

This form of identification is called a Choleski decomposition (Enders, 1995, 303). The error terms show that the contemporaneous effect of \( y_t \) on \( z_t \) being zero can be restated as the original shocks \( \epsilon_{yt} \) and \( \epsilon_{zt} \) both influencing \( y_t \), but \( z_t \) being only affected by \( \epsilon_{zt} \).

**Impulse Responses**

In order to trace the effects of shocks on the variables over time, the vector autoregressive system needs to be rewritten in its moving average form. This is accomplished in analogy to the representation of a single variable autoregression in moving average form. Take a simple AR(1)-process

\[
Y_t = \varphi Y_{t-1} + \epsilon_t
\]

If \( \varphi = 1 \), this process represents a random walk and thus is non-stationary. However, the first difference of the random walk is stationary since \( \epsilon_t \) is stationary. We thus call the random walk integrated of order one (I(1)). In order to test for stationarity, so-called ‘unit roots tests’ test for this very characteristic. If \( \varphi < 1 \), the series is stationary, in case it is above one, it would explode. We assume that \( Y_t \) is stationary and can then express it as a moving average process.

\[
Y_{t-1} = \varphi Y_{t-2} + \epsilon_{t-1} \Rightarrow Y_t = \varphi^2 Y_{t-2} + \varphi \epsilon_{t-1} + \epsilon_t
\]
\[ Y_{t-2} = \varphi Y_{t-3} + \epsilon_{t-2} \Rightarrow Y_t = \varphi^3 Y_{t-3} + \varphi^2 \epsilon_{t-2} + \varphi \epsilon_{t-1} + \epsilon_t \]

Since \( \varphi < 1 \),

\[ Y_t = \epsilon_t + \varphi \epsilon_{t-1} + \varphi^2 \epsilon_{t-2} + \varphi^3 \epsilon_{t-3} + \ldots + \varphi^q \epsilon_{t-q} \]

The moving average representation of the VAR also requires stationarity. We use matrix notation as above and have

\[ X_t = A_0 + A_1 x_{t-1} + e_t \]

Recursive replacement yields

\[ x_t = A_0 + A_1 (A_0 + A_1 x_{t-2} + e_{t-1}) + e_t = (I + A_1)A_0 + A_1^2 x_{t-2} + A_1 e_{t-1} + e_t \]

and, after replacing \( n \) times

\[ x_t = (I + A_1 + A_1^2 + \ldots + A_1^n)A_0 + A_1^{n+1} x_{t-(n+1)} + \sum_{i=0}^{n} A_1^i e_{t-n} \]

In case of stability (\( A_1^n \) disappears as \( n \) grows towards infinity) the above equation can be rewritten as

\[ x_t = \mu + \sum A_1^i e_{t-i} \quad (5.12) \]

where \( \mu = [\bar{y} \ z]^t \). Equivalently

\[ \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \]

where

\[ \bar{y} = \frac{[a_{10}(1 - a_{22}) + a_{12}a_{20}]}{(1 - a_{11})(1 - a_{22}) - a_{12}a_{21}}, \quad \bar{z} = \frac{[a_{20}(1 - a_{11}) + a_{21}a_{10}]}{(1 - a_{11})(1 - a_{22}) - a_{12}a_{21}} \]

The current value of vector \( x_t \) is expressed as a sequence of error terms \( e_{it} \), whose effect on \( x_t \) can now be traced through time. Note however that these again are the composed errors and not the structural shocks \( \epsilon_{it} \). To identify the effect of a structural shock, first rewrite 5.6 and 5.7 in matrix form.

135
\[
\begin{bmatrix}
  e_{1t} \\
  e_{2t}
\end{bmatrix} = \frac{1}{1 - b_{12}b_{21}}
\begin{bmatrix}
  1 & -b_{12} \\
  -b_{21} & 1
\end{bmatrix}
\begin{bmatrix}
  \epsilon_{yt} \\
  \epsilon_{zt}
\end{bmatrix}
\]

Replace the vector \( e_{it} \) in 5.12 by the above expression to get

\[
\begin{bmatrix}
  y_t \\
  z_t
\end{bmatrix} = \begin{bmatrix}
  \bar{y} \\
  \bar{z}
\end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum a_{11} a_{12} \begin{bmatrix}
  1 & -b_{12} \\
  -b_{21} & 1
\end{bmatrix}^i \begin{bmatrix}
  \epsilon_{yt-i} \\
  \epsilon_{zt-i}
\end{bmatrix}
\]

In order to simplify notation, define matrix \( \phi_i \) as

\[
\phi_i = \frac{A^i}{1 - b_{12}b_{21}}
\]

with elements \( \phi_{jk}(i) \). Thus, the VAR can be rewritten in a moving average form of the original shocks

\[
\begin{bmatrix}
  y_t \\
  z_t
\end{bmatrix} = \begin{bmatrix}
  \bar{y} \\
  \bar{z}
\end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix}
  \phi_{11}(i) & \phi_{12}(i) \\
  \phi_{21}(i) & \phi_{22}(i)
\end{bmatrix} \begin{bmatrix}
  \epsilon_{yt-i} \\
  \epsilon_{zt-i}
\end{bmatrix}
\]

(5.13)

The coefficients \( \phi_{jk}(i) \) can now be interpreted as impulse responses. For example, \( \phi_{12}(0) \) is the immediate impact of a shock in \( \epsilon_{zt} \) on \( y_t \)—the so-called impact multiplier. Similarly, \( \phi_{22}(1) \) represents the impact of a shock in \( \epsilon_{zt-1} \) on \( z_t \). They can be plotted against \( i \) to yield the impulse response functions, or added up to yield the cumulated impulse responses. For illustration, take a single variable AR(1) process of the form \( Y_t = 0, 6Y_{t-1} + \epsilon_t \). A change of \( \epsilon_t \) of one unit changes \( Y_t \) by one unit. In the subsequent period, the change in \( Y_{t+1} \) is 0.6, in period two 0.6² and so forth. The impulse response function in this simple case thus would be \( \phi(i) = (0.6)^i \).

Yet, the crucial problem of identification remains to be solved. One possible solution that was explained above is to set one of the parameters expressing the contemporaneous effect of one variable on the other equal to zero. Sticking with the assumption taken above, \( b_{21} = 0 \) (\( y_t \) has no contemporaneous effect on \( z_t \)), the error terms reduce to 5.10 and 5.11. This method is also called “ordering of the variables” (Enders, 1995, 307) because a shock in \( z_t \) affects both series, while a shock in \( y_t \) does not instantly impact on \( z_t \). \( z_t \) is thus called ‘prior’ to \( y_t \). The observed errors \( \epsilon_{zt} \) can be attributed fully to the structural shock \( \epsilon_{zt} \), and the sequence \( \epsilon_{yt} \) can be calculated with the help of 5.10.
Clearly, the ordering of variables is very important in this procedure and runs counter to the original advantage of VARs—that they do not require prior restrictions of the system. Ideally, the decision of how to order variables is motivated by economic theory, yet that is not always possible. Not surprisingly, this aspect is one of the major points of criticism of VARs voiced by a large number of authors (Kennedy, 2003, 347). Structural VARs as presented in the next section attempt to address this issue.

**Structural VAR analysis**

With little or no theoretical input, one has to be careful in interpreting the results of VARs identified in the way described above. It is useful in describing relationships between selected variables and in forecasting, but little can be said about causalities.

As a remedy, some authors have proposed identification procedures that are motivated explicitly by economic theory. Sims, in estimating a six-variable VAR, proposes coefficient restrictions in the matrix B that are consistent with economic theory (Enders, 1995, 330). Blanchard and Quah (1989) on the other hand restrict the cumulative effect of a demand shock, setting the respective $\sum \phi = 0$. The underlying assumption is that only a supply shock has a permanent effect on output. In the following section, linkages and sectoral development will be analyzed using another form of structural VAR analysis as proposed by Abeysinghe and Forbes (2005).

**5.2.2 Linkages and Sectoral Development**

As motivated above, I seek to investigate whether linkages between economic sectors are present and how large their contribution to individual sectoral development is. Methodologically I try to solve this puzzle by using an SVAR approach which not only allows to consider the direct input-output linkages between individual sectors, but also indirect effects that are not captured in IO-tables.¹

**The Econometric Model**

The model is based on a structural VAR model developed by Abeysinghe and Forbes (2005), which they used for the analysis of trade linkages and

¹This section is based upon joint work with Michael Wild (Schwank and Wild 2008)
output multiplier effects between countries. In our model the trade linkages are replaced by the linkages between sectors—stemming from intermediate inputs used in a sector that come from the rest of the economy. The reduced form representation of the output linkages is derived as follows.

Let $Y$ be an individual sector’s production, which is a composite of

$$Y = IM + A \quad (5.14)$$

where $IM$ is the sector’s output used as input in other sectors (intermediate output) and $A$ is the part of output that is going to final consumption, exports and that is used in its own production. Rewriting equation (5.14) in a way where $IM$ represents the sum of all the intermediate output going to the other $(n-1)$ sectors gives

$$Y = \sum_{j=1}^{n-1} IM_j + A \quad (5.15)$$

We rewrite equation (5.15) in terms of growth rates to get

$$\frac{dY}{Y} = \frac{1}{Y} \left[ \sum_{j=1}^{n-1} dIM_j + dA \right] \quad (5.16)$$

In a next step, we assume that the intermediate output to sector $j$ can be formulated as a function of total output in sector $j$.

$$IM_j = IM_j(Y_j) \quad (5.17)$$

Differentiating (5.17) with respect to $Y_j$ gives

$$dIM_j = (\frac{\partial IM_j}{\partial Y_j})dY_j \quad (5.18)$$

By inserting (5.18) into (5.16) we finally get

$$\frac{dY}{Y} = \frac{IM}{Y} \sum_{j=1}^{n-1} [\eta_j(IM_j/IM)(dY_j/Y_j)] + dA/Y \quad (5.19)$$

where $\eta_j = (\frac{\partial IM_j}{\partial Y_j})(Y_j/IM_j)$ represents the elasticity of intermediate inputs flowing from the sector under scrutiny to sector $j$ with respect to sector $j$’s production. It is thus a measure of how strongly the intermediate output flowing to a specific sector will react to this sector’s growth. By
assuming now that this elasticity is the same across sectors and adding time and industry subscripts we can then rewrite and simplify equation 5.19 as

\[ y_{it} = \alpha_i y^{im}_{it} + u_{it} \] (5.20)

where \( \alpha = (IM/Y)\eta \) and \( y^{im} = \sum (IM_j/IM)y_j \). The elasticity \( \eta \), assumed to be the same across sectors, can be taken out of the sum, which leaves other sectors’s growth rates weighted by their contribution to the total intermediate output of the sector under scrutiny. The error term \( u_{it} \) captures omitted variables.

The core proposition of this formulation is that sectoral growth depends on the growth performance of other sectors in the economy. The extent of direct influence is determined by the intermediate output linkage \( IM_j \). Yet, the model also captures indirect effects. For example, if output growth in the furniture manufacturing sector slows down, this obviously has a direct effect on the demand for machinery. There is also an indirect effect as the demand for transportation also diminishes, and thus transportation’s demand of machinery. All the other determinants of sectoral growth will show up in the error term.

In a next step, and to amend our equations to econometric estimation, we transform the model and rewrite it as a structural vector autoregression. We thus assume that sectoral output performance can be explained by its own past output performance and the past output performance of all the other sectors in the economy. The latter enter weighted according to the linkage strength. This representation closely matches equation 5.20 where sectoral growth is explained by growth in the rest of the economy, weighted by the respective linkage strength. The transformation to an autoregressive process requires the assumption that the error terms \( u_{it} \) are correlated over time and across equations and that they follow a vector ARMA process as described in Abeysinghe and Forbes (2005, 359). Formally,

\[ y_{it} = \lambda_i + \sum_{j=1}^{p} \phi_{ji} y_{it-j} + \sum_{j=0}^{p} \beta_{ji} y_{it-j} + \epsilon_{it} \] (5.21)

where \( \lambda_i \) is a vector of constants, the first sum captures all sectors’ own lagged values and the second present and past values of the other sectors weighted by linkages. More precisely,
\[ y_{it}^f = \sum_{j=1}^{n} w_{ij} y_{jt} \]

The elements of the weighting matrix, \( w_{ij} \), are the shares of total intermediate output of sector \( i \) that go to the other sectors \( j \) \((IM_j / IM)\). \( i \neq j \). The sum of all shares of sector \( i \)'s output to the other sectors must equal one. Equation (5.21) can be estimated separately for each sector of the economy. However, they can also be written as an SVAR system of equations. To illustrate, assume that there are three sectors in the economy, and the estimation takes only one lag into consideration. Then, the SVAR system of equations equals

\[
\begin{bmatrix}
1 & -\beta_{01} w_{12} & -\beta_{01} w_{13} \\
-\beta_{02} w_{21} & 1 & -\beta_{02} w_{23} \\
-\beta_{03} w_{31} & -\beta_{03} w_{32} & 1
\end{bmatrix}
\begin{bmatrix}
y_{1t} \\
y_{2t} \\
y_{3t}
\end{bmatrix}
= 
\begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\lambda_3
\end{bmatrix} +
\begin{bmatrix}
\phi_{11} & \beta_{11} w_{12} & \beta_{11} w_{13} \\
\beta_{12} w_{21} & \phi_{22} & \beta_{12} w_{23} \\
\beta_{13} w_{31} & \beta_{13} w_{32} & \phi_{33}
\end{bmatrix}
\begin{bmatrix}
y_{1t-1} \\
y_{2t-1} \\
y_{3t-1}
\end{bmatrix}
+ 
\begin{bmatrix}
\epsilon_{1t} \\
\epsilon_{2t} \\
\epsilon_{3t}
\end{bmatrix}
\]

or, in a compact form,

\[
(B_0 \cdot W)y_t = \lambda + (B_1 \cdot W)y_{t-1} + \epsilon_t
\]

where \(( \cdot \cdot \cdot )\) stands for the element-products of the matrices. The matrix on the left hand side captures contemporaneous effects of the time series on each other, the 3x3 matrix on the right hand side captures the effect of own lagged values \((\phi_{ii})\) and lagged values of the time series on each other \((\beta_{1i})\). The major difference of this approach to other SVAR works as stressed by Abeysinghe and Forbes (2005, 360) is that the model is overidentified as the elements of the weighting matrix \( w_{ij} \) are known, and that we do not have to assume that the error terms \( \epsilon_{it} \) are uncorrelated as is usually the case, but can explicitly test for it.

**Data Description**

The model is applied to a number of settings at different levels of aggregation. A list of these sectors and the abbreviations used henceforth can be found in Table 5.1.
Our data set stems from the Quantec database, a commercial database which collects economic data for South Africa from different sources. The time period covered is 1970 to 2007, so we have 37 observations per time series. A highly aggregated sector composition—we look at the development of output in agriculture, mining, the secondary and the tertiary sector—constitutes our first layer of analysis. Unit roots tests confirm their presence, thus output enters in log first differences in the estimation. In order to assess the relevance of the MEC concept for sectoral development, we use two specifications, where the second divides the secondary sector in MEC activities (excluding mining) and non-MEC manufacturing. Figure 1 in the appendix shows the development of total output in these 5 sectors.

The corresponding weighting matrices that we need in the VAR are calculated by using data on the intermediate input used in each sector stemming from the rest of the economy. Crucially, we only consider domestic intermediate inputs, because only they will induce a backward linkage locally. This differentiation between domestic and total linkage effects is not always made in related analysis, perhaps due to a lack of data. Again, the Quantec database provides this information. Although we could have used only one weighting matrix for the whole period, we decided taking the accurate weighting matrix for each year, as this captures the change in input linkages over time. An exemplary matrix that indicates linkages between sectors is presented in Section 5.2.2—note that, in line with the model specification, the intermediate input stemming from the sector itself has been set to zero. Upon inspection, we see that the secondary sector is a large consumer of intermediate goods from other sectors, thus potentially and not surprisingly the sector with the most pulling power in the economy.

The second data set consists of a sample of output development and input linkages from 1970 to 2007 for the nine 2-digit sectors of the economy and is also taken from the Quantec database. Figure 2 in the appendix shows the individual sector’s development over the whole period in absolute values. Again, we performed a unit root test in the level values and unsurprisingly detected it in almost every sector. Due to this and also to reflect the model’s formulation in growth rates, the data is transformed into log first differences as well.

Lastly, we look at 10 manufacturing subsectors, applying the same procedure. Figure 3 in the appendix displays their output performance over the period under scrutiny. The outlier in sectoral development is the Petroleum, Chemicals and Plastics sector. It performed strongly throughout the last 37
Sectoral Abbreviations

1-digit Sectors

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>Agriculture, forestry and fishing</td>
</tr>
<tr>
<td>Min</td>
<td>Mining and quarrying</td>
</tr>
<tr>
<td>MEC</td>
<td>Pet, NMM, Basic Metals, Electricity</td>
</tr>
<tr>
<td>Sec</td>
<td>Secondary Sector (Man, Elec, Cons)*</td>
</tr>
<tr>
<td>Tert</td>
<td>All Services</td>
</tr>
</tbody>
</table>

2-digit Sectors

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>Agriculture, forestry and fishing</td>
</tr>
<tr>
<td>Min</td>
<td>Mining and quarrying</td>
</tr>
<tr>
<td>Man</td>
<td>Manufacturing*</td>
</tr>
<tr>
<td>Elec</td>
<td>Electricity, gas and water</td>
</tr>
<tr>
<td>Cons</td>
<td>Construction</td>
</tr>
<tr>
<td>Trad</td>
<td>Trade, catering and accommodation services</td>
</tr>
<tr>
<td>TrSC</td>
<td>Transport, storage and communication</td>
</tr>
<tr>
<td>FinI</td>
<td>Financial intermediation, insurance, real estate</td>
</tr>
<tr>
<td>Comm</td>
<td>Community, social and personal services</td>
</tr>
</tbody>
</table>

3-digit sectors

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>Food, beverages and tobacco</td>
</tr>
<tr>
<td>Tex</td>
<td>Textiles, clothing and leather</td>
</tr>
<tr>
<td>Wood</td>
<td>Wood and paper; publishing and printing</td>
</tr>
<tr>
<td>Pet</td>
<td>Petroleum products, chemicals, rubber and plastic</td>
</tr>
<tr>
<td>NMM</td>
<td>Other non-metallic mineral products</td>
</tr>
<tr>
<td>Met</td>
<td>Metals, metal products, machinery and equipment</td>
</tr>
<tr>
<td>Elm</td>
<td>Electrical machinery and apparatus</td>
</tr>
<tr>
<td>Rad</td>
<td>Radio, TV, instruments, watches and clocks</td>
</tr>
<tr>
<td>Trans</td>
<td>Transport equipment</td>
</tr>
<tr>
<td>Furn</td>
<td>Furniture and other manufacturing</td>
</tr>
</tbody>
</table>

Table 5.1: Sectoral abbreviations, *when MEC is included, Sec and Man represent non-MEC manufacturing

years and did particularly well in the second half of the 1990s. Another important sector that grew strongly since the democratic opening is the Transport Equipment sector, reflecting the success of the industrial policy intervention in this specific area, the Motor Industry Development Programme.

In terms of linkages, there is no one sector within manufacturing that possesses a dominance equivalent to the manufacturing sector as a whole on a higher level. Overall, Transport Equipment draws the largest amount of inputs from other sectors—in accordance with the common understanding
that automotive production is desirable because of its widespread linkages and thus justifying the support the sector receives.

**Estimation Results**

The estimation follows the approach by Zellner and Palm (1974) which estimates the model in the tradition of a seemingly unrelated regressions (SUR) model instead of using the standard SVAR approach. Thus, equation 5.21 is estimated for each sector separately. It can then be rewritten and expressed as a SVAR system so that impulse responses can be calculated.

Due to the fact that we only have yearly observations, we can include only one lag of each sector. We estimated the model in OLS, 2SLS and 3SLS. For the latter two approaches the residual correlations have been calculated and the Breusch-Pagan test on the diagonality of the residual-correlation matrix has been applied. Even though this test rejects the null of diagonality for some of the specifications (those involving 9 and 10 sectors), we decided to use the 3SLS estimates for our further analysis, as the number of significant correlation coefficients is lowest when compared to the 2SLS or OLS estimates. Tables 5.2, 5.3 and 5.4 display these coefficients for the three data sets, in the specification separating MEC activities from the rest of manufacturing. The other matrices are available on request, the number of significant non-zero correlations does not change, however. Those error term correlations that are significant at the 5% are printed in bold letters. For the highest level of aggregation there are no correlations significantly different from zero, indicating that the model is correctly specified. In the case of 9 (10 when one considers MEC manufacturing separately) sectors there are only 2. However, the 10 manufacturing subsectors include 9 statistically significant correlations, two of which are negative which is certainly surprising. This might be due to poor data quality or improper specification, it definitely implies that results must be interpreted very carefully for this level of aggregation because there are unresolved problems in the estimation.

**Impulse Response Analysis**

To answer the question about the direct and indirect impact of a change in one sector on the other sectors we use impulse response analysis (Hamilton, 1994). To this end we use the model estimated above and calculate the recursive relationships according to the VAR-results. This allows us to
Table 5.2: Correlation for 3SLS

<table>
<thead>
<tr>
<th></th>
<th>Agri</th>
<th>Min</th>
<th>MEC</th>
<th>Man</th>
<th>Elec</th>
<th>Cons</th>
<th>Trade</th>
<th>TrSC</th>
<th>Finl</th>
<th>Comm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>-0.19</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEC</td>
<td>-0.09</td>
<td>0.03</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>-0.18</td>
<td>-0.22</td>
<td>0.23</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elec</td>
<td>0.39</td>
<td>-0.15</td>
<td>0.07</td>
<td>-0.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cons</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-0.04</td>
<td>0.14</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>0.19</td>
<td>0.06</td>
<td>0.14</td>
<td>0.07</td>
<td>0.16</td>
<td>-0.03</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TrSC</td>
<td>0.28</td>
<td>-0.15</td>
<td>0.48</td>
<td>-0.07</td>
<td>0.27</td>
<td>-0.09</td>
<td>0.01</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finl</td>
<td>0.13</td>
<td>0.19</td>
<td>0.05</td>
<td>-0.25</td>
<td>-0.23</td>
<td>0.01</td>
<td>0.07</td>
<td>-0.03</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Comm</td>
<td>0.03</td>
<td>0.28</td>
<td>0.09</td>
<td>-0.23</td>
<td>0.22</td>
<td>-0.16</td>
<td>0.05</td>
<td>0.11</td>
<td>0.26</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 5.3: Correlation for 3SLS, bold indicates significance at the 5 % level

isolate the effect of a unit shock in one sector on all the other sectors in the model. Technically, and as explained above in the theoretical section, the autoregressive system has to be transformed into a moving average form. The procedure is equivalent to the exposition in the theoretical section on impulse responses; starting with (5.22), by recursive replacement, we arrive at

\[ y_t = \sum_{i=0}^{\infty} C_1^i u_{t-i} \]  \hspace{1cm} (5.23)

where \( u_{t-i} = (B_0 \cdot W)^{-1} \epsilon_{t-i} \) and \( C_1 = (B_0 \cdot W)^{-1}(B_1 \cdot W) \). The impulse response matrix thus is \( C_1^i(B_0 \cdot W)^{-1} \).

Looking at the overall economy at the highest level of aggregation first, the change of sectoral shares between 1970 and 2007 reveals that the tertiary sector has gained in importance, mostly at the expense of the primary sector, while the secondary sector remained roughly stable. A closer look, distinguishing between MEC and non-MEC manufacturing, reveals however that the latter declined in importance quite substantially, while MEC activities became more dominant (see Figure 5.1). Any statement about the
Table 5.4: Correlation for 3SLS, bold indicates significance at the 5% level

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Tex</th>
<th>Wood</th>
<th>Pet</th>
<th>NMM</th>
<th>Met</th>
<th>ElecM</th>
<th>Rad</th>
<th>Trans</th>
<th>Furn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tex</td>
<td>-0.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>-0.21</td>
<td>0.31</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pet</td>
<td>-0.37</td>
<td>0.18</td>
<td>0.05</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMM</td>
<td>-0.25</td>
<td>0.17</td>
<td>0.41</td>
<td>-0.18</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>0.03</td>
<td>0.49</td>
<td>0.08</td>
<td>-0.04</td>
<td>0.16</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ElecM</td>
<td>0.42</td>
<td>0.34</td>
<td>0.04</td>
<td>0.15</td>
<td>-0.29</td>
<td>0.12</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rad</td>
<td>0.42</td>
<td>0.25</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.06</td>
<td>-0.04</td>
<td>0.62</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trans</td>
<td>-0.09</td>
<td>-0.38</td>
<td>0.31</td>
<td>-0.02</td>
<td>0.32</td>
<td>-0.22</td>
<td>-0.19</td>
<td>-0.23</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Furn</td>
<td>0.20</td>
<td>-0.03</td>
<td>-0.09</td>
<td>0.12</td>
<td>0.14</td>
<td>-0.21</td>
<td>0.08</td>
<td>0.12</td>
<td>0.34</td>
<td>1.00</td>
</tr>
</tbody>
</table>

declining dependence of South Africa on its mining sector must therefore be qualified.

![Pie charts showing sectoral output in 1970 and 2007](image)

Figure 5.1: Share of Sectoral Output in Total Output in 1970 and 2007, Source: Quantec

On the other hand, a look at the weighting matrices used in the estimation (Table 5.5 shows an exemplary matrix for the year 2007) highlights the nonetheless central role of the secondary sector in the economy. The presentation of linkages is slightly unusual, so it requires explanation. Usually, the intermediate inputs used in a certain industry are displayed in rows,
Table 5.5: Share of intermediate inputs stemming from sectors (column) used in sector (row), 2007. Source: Quantec

<table>
<thead>
<tr>
<th></th>
<th>Agri</th>
<th>Min</th>
<th>Sec</th>
<th>Tert</th>
<th></th>
<th>Agri</th>
<th>Min</th>
<th>MEC</th>
<th>Sec</th>
<th>Tert</th>
</tr>
</thead>
<tbody>
<tr>
<td>used in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>used in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediates from</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intermediates stemming from</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agri</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
<td>0.06</td>
<td>Agri</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>Min</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
<td>0.15</td>
<td>Min</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.06</td>
<td>0.17</td>
</tr>
<tr>
<td>Sec</td>
<td>0.97</td>
<td>0.96</td>
<td>0.00</td>
<td>0.79</td>
<td>MEC</td>
<td>0.02</td>
<td>0.79</td>
<td>0.00</td>
<td>0.13</td>
<td>0.29</td>
</tr>
<tr>
<td>Tert</td>
<td>0.03</td>
<td>0.04</td>
<td>0.82</td>
<td>0.00</td>
<td>Sec</td>
<td>0.96</td>
<td>0.16</td>
<td>0.49</td>
<td>0.00</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tert</td>
<td>0.03</td>
<td>0.05</td>
<td>0.39</td>
<td>0.73</td>
<td>0.00</td>
</tr>
</tbody>
</table>

with each industry represented in a column. Here, the matrix has been transposed (for reasons that will hopefully become clear immediately), so each sector and the input it uses are to be found in the respective rows. The columns thus stand for the intermediate inputs domestically produced in the agricultural sector, and flowing to the rest of the economy—the intermediate input that remains within the sector has been set to zero in accordance with the theoretical model. These input flows to other sectors are reported as shares of total intermediate input goods produced in the respective sector. Therefore, they represent backward linkages—an expansion in a sector will increase its demand for intermediates, and thus induce expansion of production in upstream sectors.

In terms of the sectors analysed here, the secondary sector is by far the largest consumer of intermediate inputs. For example, more than 97% of agricultural goods that are used as intermediate inputs in the rest of the economy go to the secondary sector. Albeit agriculture being the strongest case, the centrality of manufacturing is also evident for the other sectors. The introduction of the MEC as a separate sector slightly changes this picture. Overall, the secondary sector, now encompassing only non-MEC manufacturing, still is the single largest demander of intermediates. Yet, as expected, and justifying this specific aggregation, intermediate mining goods are overwhelmingly used in the MEC.

Impulse responses to unit shocks in all sectors reveal a strong convergence to zero after 4 years. Table 5.6 thus displays the cumulated effects of a one-unit shock in each of the sectors on the other sectors over a period of four years. The table must be interpreted as follows: the shock in a certain industry has an accumulated effect on all the other industries as displayed in the respective row. For example, in the 4-sector setting a one-unit shock in the secondary sector leads to a 0.338 unit increase of output in agriculture.
and relatively strong movements in other sectors of the economy as well—a result in line with the strong linkages detected in the manufacturing sector. The impacts of unit shocks in agriculture and mining are negligible, in the case of mining they are even negative. Lastly, spillovers from the services sector are smaller than those from manufacturing, but they are significant nonetheless, particularly when compared to agriculture and mining. Again, this is in line with linkage strength.

Once the MEC is considered separately, it becomes clear that the expansionary spillovers we observe from unit shocks in the secondary sector are mostly due to non-MEC manufacturing. An expansion of the MEC has a positive spillover on mining, but the effects on other sectors are comparatively small. A unit shock in MEC industries only leads to a 0.113 increase in manufacturing output in this specification, and growth spillovers to agricultural and services production are also much smaller than those observed with non-MEC manufacturing.

For a more detailed analysis, we looked at the nine (ten when MEC is stated separately) 2-digit sectors (disaggregating the secondary and the tertiary sectors). Graph 5.2 reestablishes the falling role of mining itself but the surge in MEC manufacturing activities. The growth in services was particularly strong in the transport, storage and communication, and in the financial intermediation, insurance and real estate sectors.

The corresponding weighting matrices are reported in the appendix (see Tables 3 and 4). With regards to the impulse responses, we can confirm

<table>
<thead>
<tr>
<th>Shock in</th>
<th>Agri</th>
<th>Min</th>
<th>Sec</th>
<th>Tert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>0.742</td>
<td>0.033</td>
<td>0.056</td>
<td>0.070</td>
</tr>
<tr>
<td>Min</td>
<td>-0.040</td>
<td>1.198</td>
<td>-0.121</td>
<td>-0.258</td>
</tr>
<tr>
<td>Sec</td>
<td>0.338</td>
<td>0.592</td>
<td>1.049</td>
<td>0.505</td>
</tr>
<tr>
<td>Tert</td>
<td>0.085</td>
<td>0.153</td>
<td>0.247</td>
<td>1.627</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shock in</th>
<th>Agri</th>
<th>Min</th>
<th>MEC</th>
<th>Sec</th>
<th>Tert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>0.721</td>
<td>0.026</td>
<td>0.040</td>
<td>0.041</td>
<td>0.047</td>
</tr>
<tr>
<td>Min</td>
<td>-0.029</td>
<td>1.327</td>
<td>-0.070</td>
<td>-0.084</td>
<td>-0.156</td>
</tr>
<tr>
<td>MEC</td>
<td>0.044</td>
<td>0.563</td>
<td>1.101</td>
<td>0.113</td>
<td>0.194</td>
</tr>
<tr>
<td>Sec</td>
<td>0.437</td>
<td>0.355</td>
<td>0.408</td>
<td>1.330</td>
<td>0.503</td>
</tr>
<tr>
<td>Tert</td>
<td>0.151</td>
<td>0.232</td>
<td>0.307</td>
<td>0.433</td>
<td>1.592</td>
</tr>
</tbody>
</table>

Table 5.6: Cumulated impulse response after 4 periods
Figure 5.2: Share of Sectoral Output in Total Output in 1970 and 2007, Source: Quantec

Table 5.7: Cumulated impulse response after 4 periods

<table>
<thead>
<tr>
<th>Shock in</th>
<th>Agri</th>
<th>Min</th>
<th>Man</th>
<th>Elec</th>
<th>Cons</th>
<th>Trad</th>
<th>TrSC</th>
<th>FinI</th>
<th>Comm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>0.736</td>
<td>0.022</td>
<td>0.034</td>
<td>0.036</td>
<td>0.023</td>
<td>0.034</td>
<td>0.026</td>
<td>0.041</td>
<td>0.053</td>
</tr>
<tr>
<td>Min</td>
<td>0.009</td>
<td>1.240</td>
<td>0.019</td>
<td>0.039</td>
<td>0.020</td>
<td>0.020</td>
<td>0.037</td>
<td>0.025</td>
<td>0.021</td>
</tr>
<tr>
<td>Man</td>
<td>0.558</td>
<td>0.711</td>
<td>1.196</td>
<td>0.725</td>
<td>0.318</td>
<td>0.650</td>
<td>0.409</td>
<td>0.858</td>
<td>0.587</td>
</tr>
<tr>
<td>Elec</td>
<td>0.005</td>
<td>0.045</td>
<td>0.011</td>
<td>1.711</td>
<td>0.044</td>
<td>0.010</td>
<td>0.007</td>
<td>0.019</td>
<td>0.007</td>
</tr>
<tr>
<td>Cons</td>
<td>0.160</td>
<td>0.227</td>
<td>0.343</td>
<td>0.233</td>
<td>2.250</td>
<td>0.259</td>
<td>0.162</td>
<td>0.403</td>
<td>0.195</td>
</tr>
<tr>
<td>Trad</td>
<td>0.133</td>
<td>0.170</td>
<td>0.266</td>
<td>0.395</td>
<td>0.541</td>
<td>1.664</td>
<td>0.364</td>
<td>0.862</td>
<td>0.237</td>
</tr>
<tr>
<td>TrSC</td>
<td>0.132</td>
<td>0.175</td>
<td>0.278</td>
<td>0.408</td>
<td>0.289</td>
<td>0.391</td>
<td>1.420</td>
<td>0.544</td>
<td>0.229</td>
</tr>
<tr>
<td>FinI</td>
<td>0.072</td>
<td>0.106</td>
<td>0.154</td>
<td>0.190</td>
<td>0.405</td>
<td>0.196</td>
<td>0.172</td>
<td>2.656</td>
<td>0.293</td>
</tr>
<tr>
<td>Comm</td>
<td>0.007</td>
<td>0.010</td>
<td>0.014</td>
<td>0.018</td>
<td>0.025</td>
<td>0.018</td>
<td>0.013</td>
<td>0.041</td>
<td>1.408</td>
</tr>
</tbody>
</table>

the findings of the above specification at a higher level (see Tables 5.7 and 5.8). Agriculture and mining provide virtually no spillovers to the rest of the economy, while manufacturing appears to have significant pulling power. Interestingly, the construction sector has comparatively strong effects, given its size. Of the services sectors, community, social and personal services have by far the weakest effect, which is to be expected in the short run.

Once the MEC is reported separately (Table 5.8), we again find its pulling power small in comparison to the rest of the manufacturing sector (with the obvious exception of the mining industry). This is notable also
Table 5.8: Cumulated impulse response after 4 periods

<table>
<thead>
<tr>
<th>Shock in</th>
<th>Agri</th>
<th>Min</th>
<th>MEC</th>
<th>Man</th>
<th>Elec</th>
<th>Cons</th>
<th>Trad</th>
<th>TrSC</th>
<th>Finl</th>
<th>Comm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agri</td>
<td>0.737</td>
<td>0.031</td>
<td>0.037</td>
<td>0.040</td>
<td>0.049</td>
<td>0.039</td>
<td>0.043</td>
<td>0.032</td>
<td>0.062</td>
<td>0.056</td>
</tr>
<tr>
<td>Min</td>
<td>0.007</td>
<td>1.223</td>
<td>0.011</td>
<td>0.012</td>
<td>0.024</td>
<td>0.014</td>
<td>0.013</td>
<td>0.020</td>
<td>0.017</td>
<td>0.013</td>
</tr>
<tr>
<td>MEC</td>
<td>0.163</td>
<td>0.728</td>
<td>1.111</td>
<td>0.279</td>
<td>0.637</td>
<td>0.315</td>
<td>0.362</td>
<td>0.308</td>
<td>0.611</td>
<td>0.343</td>
</tr>
<tr>
<td>Man</td>
<td>0.822</td>
<td>0.629</td>
<td>0.651</td>
<td>1.522</td>
<td>0.834</td>
<td>0.612</td>
<td>0.882</td>
<td>0.487</td>
<td>1.224</td>
<td>0.709</td>
</tr>
<tr>
<td>Elec</td>
<td>0.028</td>
<td>0.190</td>
<td>0.043</td>
<td>0.052</td>
<td>1.799</td>
<td>0.188</td>
<td>0.050</td>
<td>0.034</td>
<td>0.091</td>
<td>0.036</td>
</tr>
<tr>
<td>Cons</td>
<td>0.211</td>
<td>0.319</td>
<td>0.364</td>
<td>0.386</td>
<td>0.346</td>
<td>2.366</td>
<td>0.342</td>
<td>0.221</td>
<td>0.550</td>
<td>0.257</td>
</tr>
<tr>
<td>Trad</td>
<td>0.312</td>
<td>0.385</td>
<td>0.410</td>
<td>0.550</td>
<td>0.751</td>
<td>0.972</td>
<td>1.916</td>
<td>0.584</td>
<td>1.452</td>
<td>0.456</td>
</tr>
<tr>
<td>TrSC</td>
<td>0.265</td>
<td>0.369</td>
<td>0.412</td>
<td>0.482</td>
<td>0.717</td>
<td>0.596</td>
<td>0.649</td>
<td>1.492</td>
<td>0.976</td>
<td>0.403</td>
</tr>
<tr>
<td>Finl</td>
<td>0.150</td>
<td>0.205</td>
<td>0.211</td>
<td>0.273</td>
<td>0.337</td>
<td>0.600</td>
<td>0.326</td>
<td>0.258</td>
<td>2.793</td>
<td>0.397</td>
</tr>
<tr>
<td>Comm</td>
<td>0.047</td>
<td>0.064</td>
<td>0.069</td>
<td>0.085</td>
<td>0.108</td>
<td>0.138</td>
<td>0.103</td>
<td>0.070</td>
<td>0.206</td>
<td>1.372</td>
</tr>
</tbody>
</table>

because the MEC is a significant consumer of intermediates from a number of sectors—yet this conventional backward linkage is not reflected to the same extent in the impulse response results.

Lastly, we have also attempted to model spillovers within the subsectors of manufacturing. As reported above, a substantial part of the error correlations in this specification are non-zero, which points to a misspecification of the model. Omitted variables are a likely cause, since we limit ourselves to a subsection of the economy here. Sectoral growth in the rest of the economy should be controlled for and we intend to do so in a follow up version of the paper. Therefore the following results have to be interpreted carefully and should be regarded as a first and tentative evaluation of the manufacturing subsectors.

Sectoral output shares in 1970 and 2007 (Figure 5.3) reveals the petroleum products, chemicals and plastics sector as the star performer within manufacturing over the last 37 years. Only transport equipment has gained in importance in a comparable manner—probably reflecting to an extent the efforts of the Motor Industry Development Programme. On the downside, labour-intensive sectors such as food and beverages and textiles now represent a much smaller part of total manufacturing, and the same is true for the metal, metal products and machinery sector.

The weighting matrix, printed for 2007 in the appendix (Table 5), further corroborates the positive role the transport equipment sector is playing in the economy: it is the sector with potentially the greatest pulling power, being the most important user of intermediate inputs from a variety of sectors. Other than transport equipment, food and beverages, the petroleum
products and chemicals sector significantly use intermediates from the rest of manufacturing sectors.

Results of the estimation and the corresponding impulse responses (Table 5.9) largely confirm these results. A unit shock in the transport equipment sector induces a 0.896 unit output increase in electrical machinery production, a 0.781 increase in textiles and clothing and a 0.760 increase in the metal products and machinery sector. Other sectors are pulled along as well. In contrast to the pure input linkages, the metal products and machinery sector creates stronger spillovers, particularly when compared to the petroleum products and chemicals sector. Given that the latter belongs to the MEC, these results provide a further hint (caution is due because of the specification problems) that MEC activities provide less of a spillover to the rest of the economy (in this case manufacturing) than non-MEC activities.

Conclusion

This chapter set out with a hypothesis derived from economic theory and the actual development experience of South Africa: that linkages are important.
in the sectoral growth process of an economy. They provide opportunities for local upstream and downstream producers and thus shape the development path of an economy. In the case of South Africa, economic development was and continues to be strongly influenced by its minerals endowment. The manufacturing industry evolved around the mining sector, initially serving its needs, and even today, within manufacturing subsectors, those activities close to the natural resources of the country still play an important role.

We tried to empirically assess the strength of these linkages by estimating an SVAR model of sectoral output growth that explicitly incorporated the linkages between the various sectors in South Africa. With regards to the overall aggregations that we looked at first, the results strongly suggest that manufacturing as a whole is the sector with the greatest ‘pulling power’in the economy, thereby justifying the continued attention it receives from policy makers. This is particularly true for those manufacturing activities that are not counted among the MEC. Positive shocks in the agricultural and mining sectors (and, to a lesser extent, the MEC) lead to much smaller growth spurts in the rest of the economy, and the same is also true for tertiary sectors that are comparable in size to manufacturing.

On a higher level of disaggregation, the growth performances of the ten subsectors within manufacturing in general are less dependent on each other, reflecting the fact that linkages to sectors outside manufacturing—neglected in this setting—play an important role. Linkage effects therefore have less explanatory power. The sector that stands out is the transport equipment sector that has by far the strongest growth effects on the rest of manufacturing. Compared to simple linkage via intermediate inputs, we also observe a relatively prominent role of the metals, metal products and machinery

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Tex</th>
<th>Wood</th>
<th>Pet</th>
<th>NMM</th>
<th>Met</th>
<th>ElecM</th>
<th>Rad</th>
<th>Trans</th>
<th>Furn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>1.332</td>
<td>0.067</td>
<td>0.083</td>
<td>0.090</td>
<td>0.092</td>
<td>0.074</td>
<td>0.066</td>
<td>0.056</td>
<td>0.096</td>
<td>0.053</td>
</tr>
<tr>
<td>Tex</td>
<td>0.519</td>
<td>1.071</td>
<td>0.077</td>
<td>0.159</td>
<td>0.085</td>
<td>0.094</td>
<td>0.079</td>
<td>0.070</td>
<td>0.110</td>
<td>0.156</td>
</tr>
<tr>
<td>Wood</td>
<td>0.115</td>
<td>0.079</td>
<td>0.987</td>
<td>0.149</td>
<td>0.049</td>
<td>0.067</td>
<td>0.057</td>
<td>0.076</td>
<td>0.073</td>
<td>0.176</td>
</tr>
<tr>
<td>Pet</td>
<td>0.372</td>
<td>0.215</td>
<td>0.130</td>
<td>1.346</td>
<td>0.201</td>
<td>0.193</td>
<td>0.170</td>
<td>0.153</td>
<td>0.236</td>
<td>0.200</td>
</tr>
<tr>
<td>NMM</td>
<td>0.068</td>
<td>0.064</td>
<td>0.061</td>
<td>0.099</td>
<td>1.003</td>
<td>0.085</td>
<td>0.065</td>
<td>0.053</td>
<td>0.087</td>
<td>0.057</td>
</tr>
<tr>
<td>Met</td>
<td>0.198</td>
<td>0.231</td>
<td>0.136</td>
<td>0.331</td>
<td>0.194</td>
<td>1.223</td>
<td>0.437</td>
<td>0.312</td>
<td>0.312</td>
<td>0.251</td>
</tr>
<tr>
<td>ElecM</td>
<td>0.079</td>
<td>0.080</td>
<td>0.064</td>
<td>0.164</td>
<td>0.117</td>
<td>0.190</td>
<td>1.113</td>
<td>0.338</td>
<td>0.099</td>
<td>0.128</td>
</tr>
<tr>
<td>Rad</td>
<td>0.031</td>
<td>0.037</td>
<td>0.027</td>
<td>0.040</td>
<td>0.023</td>
<td>0.040</td>
<td>0.111</td>
<td>0.979</td>
<td>0.028</td>
<td>0.128</td>
</tr>
<tr>
<td>Trans</td>
<td>0.508</td>
<td>0.781</td>
<td>0.227</td>
<td>0.583</td>
<td>0.475</td>
<td>0.760</td>
<td>0.896</td>
<td>0.723</td>
<td>1.442</td>
<td>0.380</td>
</tr>
<tr>
<td>Furn</td>
<td>0.101</td>
<td>0.140</td>
<td>0.144</td>
<td>0.121</td>
<td>0.055</td>
<td>0.120</td>
<td>0.075</td>
<td>0.117</td>
<td>0.074</td>
<td>1.226</td>
</tr>
</tbody>
</table>

Table 5.9: Cumulated impulse response after 4 periods
sector, which displays stronger growth impacts on the rest of manufacturing than the petroleum products and chemicals sector.

The policy conclusions that can be drawn from this support a widely accepted stance in the literature and in policy circles: that South Africa will need to support its manufacturing sector to achieve higher overall growth rates, and that within manufacturing, more diversification and less reliance on capital-intensive sectors close to the mineral endowment such as chemicals is needed. It also vindicates support for the auto industry, given its central role within the manufacturing subsectors.