CENTRE FOR APPLIED MACROECONOMIC ANALYSIS

The Australian National University

CAMA Working Paper Series **August, 2008**

THE INDIAN GROWTH MIRACLE AND ENDOGENOUS GROWTH

Jakob B. Madsen
Monash University

Shishir Saxena
Monash University

James B. Ang
Monash University

CAMA Working Paper 29/2008
http://cama.anu.edu.au

THE INDIAN GROWTH MIRACLE AND ENDOGENOUS GROWTH

Jakob B. Madsen,* Shishir Saxena and James B. Ang

Department of Economics, Monash University

Abstract

Using over half a century of R&D data for India, this paper examines the extent to which India's recent growth experience can be explained by R&D, international R&D spillovers, catch-up to the technology frontier and financial liberalization. Furthermore, the paper also tests whether any of the competing second-generation endogenous growth theories can explain India's growth experience. The findings provide support for Schumpeterian growth theory and indicate that the recent high growth rates in India are likely to continue well into the future.

Keywords: Schumpeterian growth; semi-endogenous growth; R&D

JEL classification: O3; O4

---

* Corresponding author: Department of Economics, Monash University, 900 Dandenong Road, Caulfield East, Vic 3145 Australia. E-mail: Jakob.Madsen@buseco.monash.edu.au.

Acknowledgments: Valuable comments and suggestions from workshop participants at Monash University, conference participants of the Harvard Project on Asian and International Relations at Harvard University, the 13th Australasian Macroeconomics Workshop at University of Sydney and the 2008 Asia Pacific Week at the Australian National University are gratefully acknowledged. Saxena thanks financial support received from the Monash Postgraduate Publication Award.

1. Introduction

While R&D and human capital have been assumed to be the key factors driving growth in endogenous growth models, so far there has been little empirical analysis that explores the role of R&D in explaining growth for the "miracle" economies. Consequently, whether there are scale-effects in the rate of innovation and hence whether the high growth rates in the "miracle economies" will continue into the future remain unknown. In fact, the implications of R&D-based theories for developing economies remain unclear (Jones, 1995a). Using data for India for over half a century, the main contributions of this paper include: 1) examining the importance of R&D in explaining growth rates in India; and 2) testing which endogenous growth theory is most applicable in explaining growth in India. The growth theories tested in this paper include the following two most influential endogenous growth models: 1) the semi-endogenous growth theories of Jones (1995a), Kortum (1997) and Segerstrom (1998); and 2) the Schumpeterian growth theories of Aghion and Howitt (1998), Young (1998) and Howitt (1999).

This study is, to the best of our knowledge, the first attempt to test the importance of R&D for growth and the validity of second-generation models for a developing country.[1] An important question addressed in this paper is whether R&D activities play a crucial role in explaining growth in developing countries or whether R&D-driven growth is limited to highly developed countries. The analysis is limited to India given that it is one of the very few developing countries for which sufficiently long R&D data are available to enable tests of the importance of R&D for growth in general, and specifically to discriminate between the second-generation growth models. The R&D data for India cover the period from 1950 to 2005 – a time span that even exceeds the R&D data that are available for almost all OECD countries.[2] Most R&D data for developing countries are available only from the 1990s, which is far too short a period to discriminate between growth models using aggregate data.

There is another important reason why India is an ideal candidate to test the R&D-driven growth and to discriminate between different endogenous growth models. India is a relatively low income country that has gone through a period of low to high growth rates. This raises the question of whether the lack of R&D has been a key factor behind the low growth rates experienced by India. In other words, has R&D been an important factor behind India's take-off in the late 1980s? In view of the coincidence of the transition from low to high growth and significant economic reforms in several key sectors of the Indian economy since the late 1980s, the literature often argues

[1] Since the first-generation models have been refuted by the empirical evidence of Jones (1995b) for the US, the empirical testing for endogenous growth theories has recently focused on discriminating between the two competing theories of the second-generation endogenous growth models using time series data for the US or a larger sample of industrialized countries (e.g., Zachariadis, 2003, 2004; Laincz and Peretto, 2006; Ha and Howitt, 2007; Madsen, 2008).
[2] R&D data are available from 1953 for the US and Japan. For most of the remaining OECD countries R&D data are only available from the late 1960s (see Madsen, 2007b).

that the high growth rates have been driven by economic liberalization or an attitudinal change favoring the reforms (Panagariya, 2002; De Long, 2003; Rodrik and Subramanian, 2005). This opens the possibility that the increasing growth rates experienced by India since the 1980s may not have been significantly affected by R&D. Therefore, it remains to be seen whether any R&D-based endogenous theory can help explain the increasing growth in India during the post-reform period. Moreover, to date, the sources of the concomitantly rising productivity growth in India over this period remain largely unexamined (Acharya, 2004). In light of this, both aggregate data (1950-2005) and firm-level data (590 firms over the period 1993-2005) are used in this paper to give insight into the ability of the second-generation growth models in explaining productivity growth for India.

The innovation-driven endogenous growth theories and their growth implications are presented and briefly discussed in the next section. Data and graphical evidence are presented in Section 3. Integration and cointegration tests using aggregate data are performed and the results are discussed in Section 4. Section 5 provides estimates of TFP growth. In addition to discriminating between the second-generation endogenous growth models, the empirical analysis also tests for the role of international technology spillovers, distance to the frontier, and absorptive capacity in driving productivity growth. Analysis using firm-level data are undertaken in Section 6 to complement the results based on aggregate data. The last section concludes the paper.

2. R&D-based Endogenous Growth Models

The following ideas production function can be used to discriminate between endogenous growth theories (see, e.g., Ha and Howitt, 2007; Madsen, 2008):

$$g_A = \frac{\dot{A}}{A} = \lambda \left(\frac{X}{Q} \right)^{\sigma} A^{\phi - 1}, \; 0 < \sigma \leq 1, \; \phi \leq 1 \tag{1}$$

$$Q \propto L^{\beta} \text{ in steady state,}$$

where g_A is total factor productivity (henceforth TFP) growth, A is TFP, X is research input (semi-endogenous growth models) or the productivity-adjusted research input (Schumpeterian growth theory), Q is product variety, L is employment, X/Q is research intensity, λ is the R&D productivity parameter, σ is a duplication parameter, which is assumed to be zero if all innovations are replication of existing knowledge and 1 if none of the new innovations are replications, ϕ is returns to scale in knowledge production and β is the coefficient of product proliferation. The key distinction between these two endogenous growth models lies in the values of ϕ and β. Semi-endogenous growth theory predicts $\phi < 1$ and $\beta = 0$, Schumpeterian-growth theory posits $\phi = 1$ and $\beta = 1$, and first-generation fully endogenous growth models assume $\phi = 1$ and $\beta = 0$.

One common feature in these models is that growth is driven by R&D. In the first-generation growth models in which growth is proportional to the number of R&D workers, or $g_A = \lambda X^{\sigma}$, the growth rate can be enhanced by increasing the number of R&D workers in the economy. Using data for the G5 countries, Jones (1995b) dismisses the first-generation growth models by showing that the TFP growth rate has not increased in the post WWII period despite a significant increase in the number of R&D workers. To overcome the problem associated with the first-generation growth models, Jones (1995b), Kortum (1997) and Segerstrom (1998) abandon the assumption of scale effects in ideas production, and replace it with the assumption of diminishing returns to knowledge. This assumption implies that increasing the number of R&D workers will enhance the TFP growth rate only temporarily whereas the permanent growth effects are zero. Consequently, a positive growth in R&D is required to sustain a positive TFP growth.

The Schumpeterian growth models of Aghion and Howitt (1998), Howitt (1999) and Dinopoulos and Thompson (2000) retain the assumption of constant returns to knowledge from the first-generation models. Thus, the long-run policy implications of the first-generation models are applicable in these models. However, due to product proliferation, they argue that the increasing level of resources devoted to R&D spread over the concomitantly increasing product variety in the economy, which sterilizes the effects of R&D productivity. The underlying argument of these models is that as the economy progresses, the possibility of firms entering the industry with either horizontally or vertically differentiated new products also increases. Aggregate R&D expenditures and R&D workers spread over this increasing variety of products or sectors. As Laincz and Peretto (2006) point out, Schumpeterian growth model eliminates the scale effects because it focuses on the scale of the firm, and not the scale of the economy. Accordingly, these models imply that R&D intensity, defined as the aggregate resources devoted to R&D per product, is postulated to be proportionally related to TFP growth.

The log-linear approximation of Eq. (1) can be written as follows:

$$\Delta \ln A_t = \ln \lambda + \sigma \left[\ln X_t - \ln Q_t + \left(\frac{\phi - 1}{\sigma} \right) \ln A_t \right] \tag{2}$$

If $\Delta \ln A_t$ is stationary, as it should be in steady state, the square bracket term should also be stationary and thus $\ln X_t$, $\ln Q_t$ and $\ln A_t$ should form a cointegrated relationship. On the other hand, if TFP growth is not stationary, as predicted by the first-generation endogenous growth models, then the variables in the square bracket will not form a cointegrated relationship, i.e., $\ln X_t$ will be unrelated to $\ln Q_t$ and $\ln A_t$ in the long run.

Since $\Delta \ln A_t$ is stationarity, as shown below, the second-generation endogenous growth theories imply that the terms υ_t and ς_t in the following equations are stationary:

$$\upsilon_t = \ln X_t + \left(\frac{\phi - 1}{\sigma}\right) \ln A_t , \qquad \text{Semi-endogenous growth theory} \qquad (3)$$

$$\varsigma_t = \ln X_t - \ln Q_t . \qquad \text{Schumpeterian growth theory} \qquad (4)$$

These two equations are used to test whether the second-generation growth models are consistent with India's growth experience. If semi-endogenous growth theory holds, one would expect i) both $\ln X_t$ and $\ln A_t$ to be non-stationary and integrated at the same order; and ii) both variables to move closely together in the long-run with a cointegrated vector $\left(1, \frac{\phi - 1}{\sigma}\right)$ in which the second element is expected to carry a negative sign due to the assumption of diminishing returns to knowledge production, i.e., $\phi < 1$, (Ha and Howitt, 2007). These two hypotheses provide a general framework to test semi-endogenous growth theory. However, if the data are consistent with Schumpeterian growth theory, one would expect i) $\ln(X/Q)_t$ to be stationary, or *I*(0), so that measures of R&D intensity exhibit no large persistent movements; and ii) both $\ln X_t$ and $\ln Q_t$ to be cointegrated with a cointegrated vector $(1, -1)$ so that there is a one-to-one relationship between R&D inputs and measures of product variety in the long run.

Madsen (2008) argues that these cointegration tests are necessary but not sufficient conditions to test the validity of these endogenous growth theories. While the cointegration predicted by Eqs. (3) and (4) may prevail, it cannot be ruled out that cointegration may be driven by forces that are unrelated to the second-generation endogenous growth models. For example, if the governments target a constant level of research intensity, Eq. (4) may be satisfied but growth needs not be related to research intensity. Thus, the following model is also estimated (for derivation, see Madsen, 2008):

$$\Delta \ln A_t = \gamma_0 + \gamma_1 \Delta \ln X_t^d + \gamma_2 \Delta mi_t \ln X_t^f + \gamma_3 \ln\left(\frac{X}{Q}\right)_t^d + \gamma_4 mi_t \ln\left(\frac{X}{Q}\right)_t^f$$

$$+ \gamma_5 \ln\left(\frac{A^{US}}{A^{IN}}\right)_{t-1} + \gamma_6 \left(\frac{X}{Q}\right)_t^d \ln\left(\frac{A^{US}}{A^{IN}}\right)_{t-1} + \gamma_7 \ln FL_t + \gamma_8 \Delta \ln FL_t + e_t , \qquad (5)$$

where mi_t is import intensity defined as the ratio of imports to GDP, A_t^{US} is the TFP level of the US, which represents the leading edge technology, the superscripts d and f stand for domestic and foreign, $(A^{US}/A^{IN})_{t-1}$ measures the distance to frontier, $(X/Q)_t^d \ln(A^{US}/A^{IN})_{t-1}$ captures the absorptive capacity, FL_t is an index of financial liberalization and e_t is a stochastic error term. This model embeds the predictions of the open economy second-generation endogenous growth models. The distance to the frontier and its interaction with research intensity follow the predictions of the Schumpeterian models of Howitt (2000) and Griffith *et al.* (2003, 2004). The above specification also allows international technology spillovers through the channel of imports due to Coe and Helpman (1995), and Madsen (2007b, 2008). Finally, an index of financial liberalization is included in the model to control for the potential positive growth effects driven by financial sector reforms in India. This variable is included both in levels and differences to allow for the possibility that both the level and growth of financial liberalization could affect TFP growth in the long run. The construction of this index is detailed in Ang (2008a) and briefly discussed in the Appendix.

While semi-endogenous growth theory predicts $\gamma_1, \gamma_2 > 0$ and $\gamma_3, \gamma_4, \gamma_5 = 0$, Schumpeterian growth theory predicts $\gamma_3, \gamma_4, \gamma_5 > 0$ and $\gamma_1, \gamma_2 = 0$. It is expected that γ_6, γ_7 and $\gamma_8 > 0$. Financial liberalization is expected to affect growth positively because it is associated with greater mobilization of savings and more efficient allocation of resources (McKinnon, 1973; Shaw, 1973).

Knowledge spillovers through the channel of imports affect TFP growth, according to some of the endogenous growth models described in Grossman and Helpman (1991). In these models TFP depends on the horizontally and vertically differentiated intermediate inputs. For horizontally differentiated products, an increasing variety of intermediate inputs increases the economy-wide efficiency of production. For vertically differentiated products, intermediate products come in different qualities and the effectiveness of an intermediate input in final production is positively related to the number of times the input has been improved. Common in both cases is that the variety and the quality of intermediate inputs are predominantly explained by cumulative R&D and, therefore, that TFP is a positive function of cumulative R&D. This line of reasoning suggests that the TFP of a country depends on its own cumulative R&D and the cumulated R&D embodied in imported intermediate inputs and, therefore, that technology is transmitted internationally by the import-weighted stock of knowledge. Weighted foreign R&D stocks are interacted with the degree of openness (measured by the ratio of import to GDP) to adequately capture the role of international trade (Coe and Helpman, 1995).

The growth effects of the distance to the frontier follow the Schumpeterian models of Howitt (2000) and Griffith et al. (2003). They show that TFP growth depends on the distance to the frontier and its interaction with research intensity. The further a country is away from the technology frontier the better it can take advantage of the technology developed by the frontier country. Furthermore, the higher is the research intensity of a country the stronger is the capacity of the country to absorb and assimilate the technology that is developed elsewhere. However, the interaction between research intensity and the distance to the frontier will not be included in all estimation because this term almost forms a linear combination with the domestic research intensity.

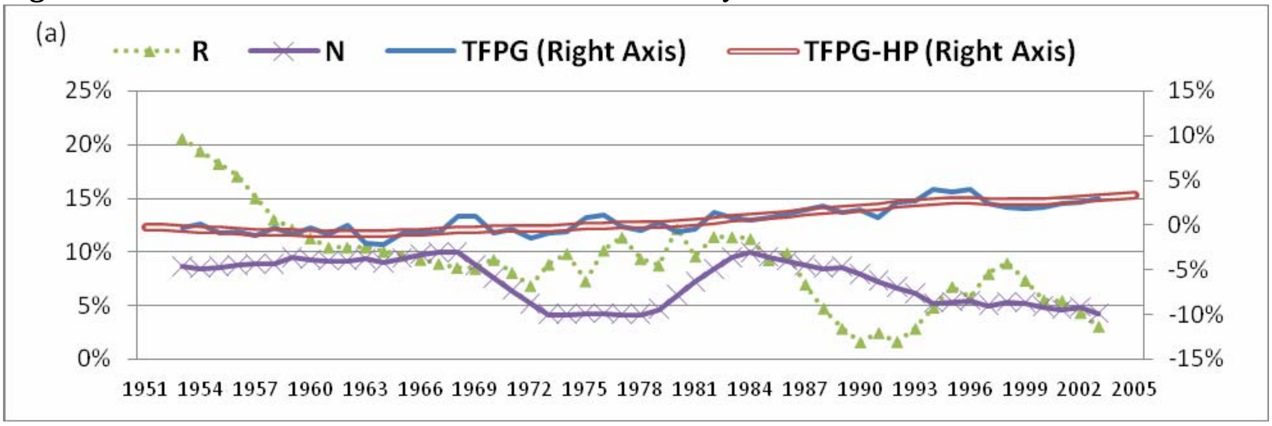
3. Data

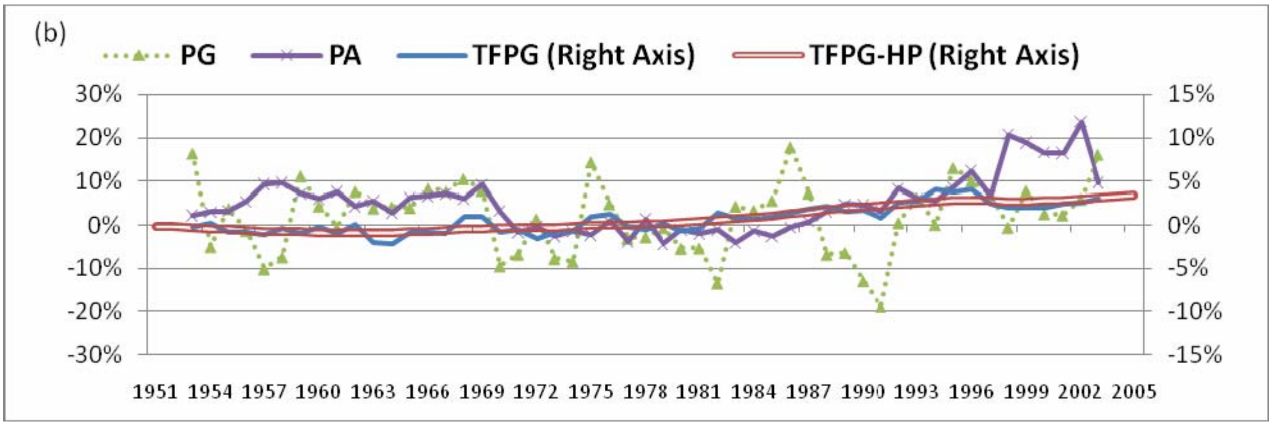
The models are estimated using aggregate data and firm level data. The analysis of the firm level data is relegated to Section 6. Annual data for the period 1950-2005 are used in the aggregate analysis. R&D input (X_t) is measured by the number of scientists and engineers engaged in R&D (N_t), real R&D expenditures (R_t), number of patent applications by domestic residents (PA_t) and the number of patents granted to domestic residents (PG_t). In the testing of Schumpeterian growth models, research intensity $(X/Q)_t$ is measured as: $(N/L)_t$, $(N/AL)_t$, $(R/AL)_t$, $(R/Y)_t$, $(PA/L)_t$ and $(PG/L)_t$, where L_t is labor force. The second and third terms are adjusted by TFP to allow for innovative activity becoming more complex as the economy advances (Aghion and Howitt, 1992). Data sources and construction of variables are explained in the Appendix.

Figures 1a and 1b compare the growth in TFP with the growth rates in R&D workers employed, real R&D expenditures, patent applications and patents granted. If the prediction of semi-endogenous growth theory is valid, one would expect TFP growth to move closely with the growth rates of R&D activity over time. It is evident that TFP growth shows a mild increasing trend since the 1970s. However, except patent applications, the growth rates of all R&D activity measures do not display an increasing trend. The growth in the number of R&D workers and real R&D expenditures has shown a downward rather than an upward trend (Figure 1a). Turning to patent measures, the growth rate of patent applications shows little co-movements with TFP growth over time whereas the growth rate in patents granted does (Figure 1b). On the whole, the graphical evidence presented here seems to lend only limited support to semi-endogenous growth theory.

Figure 2 plots various measures of research intensity and the TFP growth rate. The R&D-based measures of research intensity increased steadily from 1950 until the late 1980s or the early 1990s along with the increasing TFP growth rate. Except $(N/L)_t$ which continued to rise after 1990, there has been little variation in all these series since then. For the patent-based measures, the ratio of $(PG/L)_t$ seems to move pro-cyclically with TFP growth. Although the ratio of $(PA/L)_t$ does not seem to exhibit any strong correlation pattern with TFP growth, it has continued to rise after the reform period, reflecting the increasing drive towards innovation during this period. On the whole, the graphical evidence gives some support to Schumpeterian growth theory.

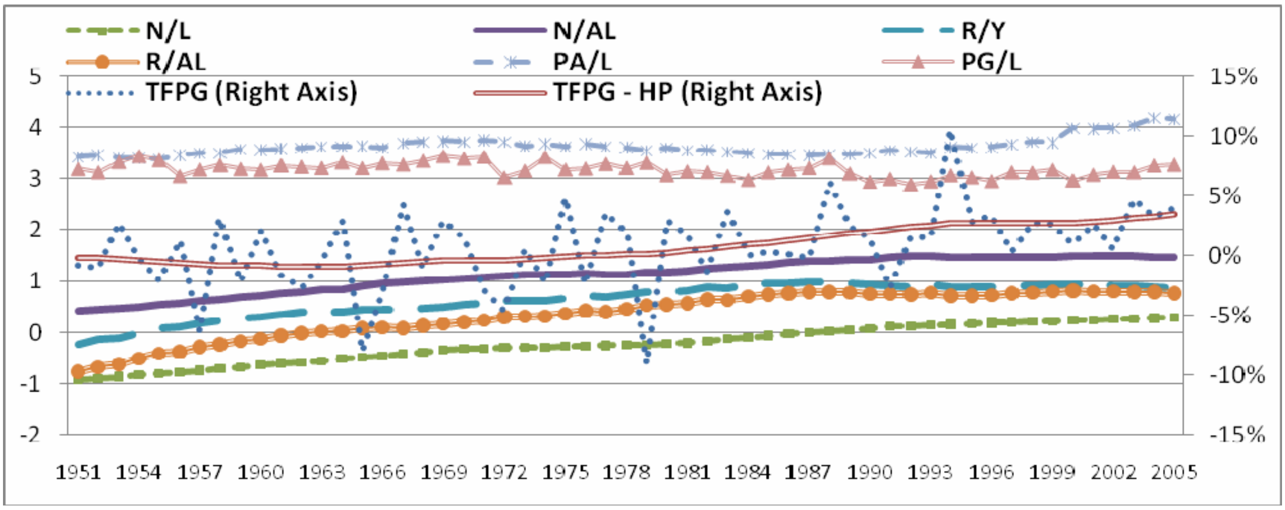
Figure 1: TFP Growth and Growth rates of R&D activity measures





Notes: *R* is real R&D expenditures; *N* is number of scientists and engineers engaged in R&D; *PA* is patent applications and *PG* is patents granted. Growth in TFP and all measures of research activity are based on 5-year centered moving average. TFPG-HP is the Hodrick-Prescott filtered series for TFP growth.

Figure 2: TFP growth and measures of research intensity



Notes: R_t is real R&D expenditures; N_t is the number of scientists and engineers engaged in R&D; A_t is total factor productivity; PA_t is patent applications; PG_t is patents granted and L_t is total employment. TFPG-HP is the Hodrick-Prescott TFP filtered series. Some of the series are scaled for the ease of comparison. All measures of research intensity are on logarithmic scales.

4. Integration and Cointegration Analyses

4.1 Unit root tests

The integration properties of macroeconomic variables are commonly examined using two standard unit root tests - the Augmented Dickey-Fuller (1979) and Phillips-Perron (1988) tests. However, these tests are known to suffer from a finite sample power and size bias, especially when the macroeconomic series is short and has structural breaks. We therefore implement the unit root tests of Ng and Perron (2001) that takes into account of the possible presence of a structural break.

Unit root tests are performed for sample periods: pre-reform (1950-1990) and full sample (1950-2005). The results reported in Table 1 indicate that the null hypothesis of a non-stationary TFP growth is rejected at the one percent level of significance regardless of the sample period considered. That is, the level of TFP (i.e., $\ln A_t$) is non-stationary in levels but it achieves stationarity after taking first difference (i.e., $\Delta\ln A_t$). Thus, while TFP level contains a unit root, TFP growth is a stationary process. This provides the basis of cointegration tests to distinguish the compatibility of Schumpeterian and semi-endogenous growth models with the Indian aggregate data. Since $\ln A_t$ follows an $I(1)$ process, one would expect R&D activity measures to contain a unit root as well to satisfy the predictions of semi-endogenous growth theory. In Table 2, unit root tests based on the Ng and Perron (2001) procedure suggest that except for $\ln PG_t$ in the pre-reform period, there is no evidence that R&D inputs follow an $I(1)$ process in either of the considered periods. From this it follows that there is little evidence in favor of semi-endogenous growth models.

Table 1: Unit root tests for TFP level and TFP growth

	Pre-reform (1950-1990)		Full sample (1950-2005)	
	$\ln A_t$	$\Delta\ln A_t$	$\ln A_t$	$\Delta\ln A_t$
Ng-Perron MZ_a^d test statistic	-6.674	-18.282***	-0.108	-26.924***
Critical values: 1%	-23.800	-13.800	-23.800	-13.800
5%	-17.300	-8.100	-17.300	-8.100
10%	-14.200	-5.700	-14.200	-5.700

Notes: the optimal lag length was selected using AIC by allowing for a maximum of nine lags. AR-GLS detrending method was used as the spectral estimation method. The reported test statistic is the modified form of Phillips-Perron test MZ_a^d. *** indicates 1% level of significance.

The unit root tests of the natural logs of $(N/L)_t$, $(N/AL)_t$, $(R/AL)_t$, $(R/Y)_t$, $(PA/L)_t$ and $(PG/L)_t$ reveal that in two out of six cases, the R&D intensity measures are stationary in levels for the pre-reform period. For the whole sample five out of six measures of research intensity are stationary. Therefore, the results seem to imply that Schumpeterian model is more relevant in explaining the TFP growth in India in the post-reform period. Overall, our results thus far suggest that the Indian data are more in line with the predictions of Schumpeterian growth theory.

Table 2: Unit root tests for semi-endogenous and Schumpeterian growth models

	Pre-reform (1950-1990)	Full sample (1950-2005)
	Support for semi-endogenous?	
$\ln N_t$	No	No
$\ln R_t$	No	No
$\ln PA_t$	No	No
$\ln PG_t$	Yes	No
	Support for Schumpeterian?	
$\ln(N/L)_t$	No	No
$\ln(N/AL)_t$	No	Yes
$\ln(R/AL)_t$	No	Yes
$\ln(R/Y)_t$	Yes	Yes
$\ln(PA/L)_t$	No	Yes
$\ln(PG/L)_t$	Yes	Yes

Notes: the optimal lag length was selected using AIC by allowing for a maximum of nine lags. AR-GLS detrended method was used as the spectral estimation method. Statistical significance at the five percent level was used as the decision rule.

4.2 Cointegration tests

4.2.1 Semi-endogenous growth theory

The Johansen approach is used to examine the existence of a long-run relationship between TFP and R&D activity (Eq. 3) as predicted by semi-endogenous growth theory. Given the sample size, we have considered a maximum lag length of four. Using the AIC criterion, the optimal lag length is found to be one for all models. Table 3 presents the results of the Johansen cointegration tests and the cointegrating vector associated with each bivariate VECM. It is evident that while cointegration is found only in two cases for the pre-reform period, both the results of the trace and maximum eigenvalue tests unanimously suggest that a cointegrated relationship exists in all cases for the whole sample period. These results are inconsistent with the unit root tests reported earlier.

Despite the findings of cointegration, a closer examination of the cointegrating vectors suggests that there is little support for semi-endogenous growth theory. In particular, in five out of eight cases, the second element in the cointegrating vectors does not carry the right sign as predicted by this theory. Although $\ln A_t$ in the remaining three cases is associated with a negative sign, it is only statistically significant in the third model for the full sample period. Furthermore, the magnitude of the coefficient also appears to be unreasonably large (-74.648). In sum, while the cointegration tests seem consistent with the predictions of semi-endogenous growth theory, the estimated cointegrating vectors are not in line with this theory, which predicts a negative and statistically significant term in the second element of the cointegrating vector. Thus, the results cast doubts on the assumption of diminishing returns to knowledge as maintained by these models.

Table 3: Johansen cointegration tests for semi-endogenous growth theory

Model	Pre-reform (1950-1990)			Full sample (1950-2005)		
	Trace statistic (Max-eigenvalue statistic)	Cointegrating vector		Trace statistic (Max-eigenvalue statistic)	Cointegrating vector	
$\ln N_t$ and $\ln A_t$	5.621 (4.722)	1.000	-12.759 [-0.992]	14.689^{*} (14.409^{**})	1.000	2.464 [1.054]
$\ln R_t$ and $\ln A_t$	16.209^{**} (15.596^{**})	1.000	2.353 [0.472]	26.604^{***} (25.943^{***})	1.000	0.478 [0.295]
$\ln PA_t$ and $\ln A_t$	11.457 (6.378)	1.000	13.197^{***} [4.029]	19.447^{**} (15.563^{**})	1.000	-74.648^{***} [-4.228]
$\ln PG_t$ and $\ln A_t$	23.697^{***} (20.778^{***})	1.000	2.619^{***} [2.952]	17.334^{**} (14.264^{*})	1.000	-0.402 [-0.735]

Notes: The null hypothesis of the test is that of no cointegrating relationship between the variables. The estimation includes an intercept but not a deterministic trend. Figures in square brackets are *t*-statistics.

5.2. Schumpeterian growth theory

For Schumpeterian growth theory, the Johansen approach is also used to examine the existence of a long-run relationship between R&D activity and product variety (Eq. 4). The results reported in Table 4 show that the null of no cointegration is rejected at conventional significance levels in most cases. The cointegration test results reported here are also more consistent with the unit root test results reported earlier. For the pre-reform period, the second elements in the cointegrating vectors are statistically significant and have the correct sign as predicted by the theory in all cases. Note that none of these estimated coefficients are close to minus one as predicted by Schumpeterian theory. However, as pointed out by Madsen (2008) this reflects the lack of a perfect representation of product variety by these variables. Therefore, this evidence is not strong enough to reject Schumpeterian growth theory. For the whole sample period, there is strong support for Schumpeterian theory in two cases. On the whole, one can conclude from these cointegration tests that the Indian aggregate data are more consistent with the predictions of Schumpeterian growth theory.

Table 4: Johansen cointegration tests for Schumpeterian growth theory

Model	Pre-reform (1950-1990) Trace statistic (Max-eigenvalue statistic)	Cointegrating vector		Full sample (1950-2005) Trace statistic (Max-eigenvalue statistic)	Cointegrating vector	
$\ln N_t$ and $\ln L_t$	13.481* (13.039*)	1.000	-3.671*** [-24.075]	20.036** (19.177***)	1.000	-3.550*** [-35.942]
$\ln N_t$ and $\ln AL_t$	6.224 (6.021)	1.000	-2.031** [-1.774]	15.494** (14.264**)	1.000	1.787 [1.561]
$\ln R_t$ and $\ln AL_t$	18.369** (17.647**)	1.000	-3.003*** [-3.988]	28.674*** (26.591***)	1.000	-0.016 [-0.026]
$\ln R_t$ and $\ln Y_t$	16.380** (15.907**)	1.000	-1.918*** [-8.634]	28.654*** (25.698***)	1.000	0.181 [0.451]
$\ln PA_t$ and $\ln L_t$	22.657*** (20.646***)	1.000	-0.732*** [-3.761]	12.711 (12.337*)	1.000	0.069 [0.137]
$\ln PG_t$ and $\ln L_t$	19.876** (19.123***)	1.000	-0.386* [-1.709]	19.271** (16.062**)	1.000	-0.436** [-2.236]

Notes: The null hypothesis of the test is that of no cointegrating relationship between the variables. The estimation includes an intercept but not a deterministic trend. Figures in square brackets are *t*-statistics.

5. TFP Growth Estimates

Following the discussion in Section 2, we provide the empirical estimates for the TFP growth equation. Eq. (5) is first estimated without considering the effects of absorptive capacity, and it is estimated in 5- and 10-year differences. Although the equation can also be estimated in 1-year differences, these regressions are not performed given that results based on 1-year differences may be affected by transitional dynamics and business cycles. Since the dependant variable in all cases is found to be stationary, the equation is estimated using ordinary least squares. The research intensity measures are taken to be the average over all years covered by the differences. Distance to the frontier and absorptive capacity are evaluated at the beginning of the period for the differences considered. In view that there is some evidence of serial correlation, we adjust the standard errors using the Newey-West procedure in order to obtain heteroskedastic and autocorrelation consistent estimates.

The results are reported in Table 5. First, consider the estimates related to semi-endogenous growth theory. The coefficient of growth in domestic research activity ($\Delta\ln X^d$) is found to be either statistically or economically insignificant in most cases, and thereby providing little evidence in support of semi-endogenous growth theory.[3] There is also no evidence to suggest that growth in imports of technology ($\Delta \ln X^f$) tends to promote TFP growth in India. Its coefficient is found to be statistically insignificant for estimates based on both 5- and 10-year differences. This finding is perhaps not surprising given the highly restrictive nature of the trade regime in India before 1991. Workers from developing economies need time and adequate skills to be able to assimilate the tacit technology embodied in these imports (Savvides and Zachariadis, 2005). If workers in these countries are sufficiently skilled, which appears to be the case for India as pointed out by Hall and Jones (1999), then utilizing the information embodied in imported technology could be difficult. Using OECD data, the findings of Madsen (2008) also do not show any significant role of growth of imported technology for both 5- and 10-year difference estimates.

Table 5: Estimates of TFP growth with five- and ten-year differences

	$X = N$ $(X/Q)^d = N/L$ $(X/Q)^f = (R/Y)^f$		$X = N$ $(X/Q)^d = N/AL$ $(X/Q)^f = (R/Y)^f$		$X = R$ $X/Q = R/Y$		$X = R$ $X/Q = R/AL$		$X=PA$ $X/Q = PA/L$		$X = PG$ $X/Q = PG/L$	
						5-year differences						
Intercept	-2.49	-(7.11)	-1.50	-(3.79)	-1.81	-(4.14)	-2.61	-(6.07)	-1.95	-(5.03)	-1.95	-(3.50)
$\Delta \ln X^d$	-0.29	-(2.98)	-0.33	-(3.29)	-0.11	-(1.65)	-0.12	-(1.84)	0.02	(1.10)	0.02	(2.50)
$\Delta mi \ln X^f$	-0.09	-(0.30)	-0.07	-(0.25)	-0.07	-(0.21)	-0.25	-(0.71)	0.48	(1.18)	0.07	(0.22)
$\ln(X/Q)^d$	0.11	(4.60)	0.14	(4.28)	0.21	(2.46)	-0.06	-(1.24)	-0.01	-(0.29)	-0.05	-(0.62)
$mi \ln(X/Q)^f$	3.07	(5.88)	3.18	(5.94)	1.61	(1.91)	2.75	(5.12)	4.94	(5.75)	-1.36	-(1.75)
$\ln DTF$	0.90	(8.55)	0.89	(8.28)	0.81	(8.36)	0.85	(7.81)	0.72	(6.41)	0.81	(8.20)
$\ln FL$	0.29	(5.50)	0.34	(5.28)	0.33	(2.97)	0.15	(1.65)	0.16	(3.60)	0.08	(0.62)
$\Delta \ln FL$	0.04	(0.77)	0.05	(0.78)	0.09	(1.07)	0.06	(1.03)	0.22	(4.17)	0.01	(0.22)
LM-1	2.41	[0.12]	2.82	[0.09]	0.18	[0.68]	0.94	[0.33]	2.27	[0.13]	0.63	[0.43]
LM-2	3.27	[0.20]	3.82	[0.15]	0.98	[0.61]	0.94	[0.62]	4.18	[0.12]	0.75	[0.69]
Normality	0.08	[0.95]	0.08	[0.96]	0.92	[0.62]	0.48	[0.78]	1.51	[0.46]	1.75	[0.41]
						10-year differences						
Intercept	-2.26	-(6.77)	-0.54	-(1.33)	-1.65	-(2.53)	-2.80	-(5.16)	-1.96	-(3.75)	324.39	(0.00)
$\Delta \ln X^d$	-0.19	-(2.64)	-0.23	-(3.24)	-0.17	-(2.63)	-0.17	-(4.61)	0.08	(2.48)	0.02	(0.94)
$\Delta mi \ln X^f$	-0.07	-(0.31)	-0.01	-(0.04)	0.13	(0.59)	-0.10	-(0.41)	0.37	(0.83)	-0.33	-(1.09)
$\ln(X/Q)^d$	0.19	(7.76)	0.25	(7.58)	0.49	(2.91)	-0.11	-(1.06)	-0.06	-(1.97)	0.12	(0.59)
$mi \ln(X/Q)^f$	4.77	(5.83)	4.96	(6.29)	1.52	(0.95)	4.35	(3.75)	6.69	(2.98)	-3.80	-(2.31)
$\ln DTF$	0.87	(9.95)	0.85	(9.88)	0.88	(7.30)	0.92	(9.51)	0.84	(6.28)	0.78	(6.84)
$\ln FL$	0.32	(6.04)	0.40	(6.74)	0.67	(2.98)	0.11	(0.59)	0.12	(1.61)	0.22	(0.59)
$\Delta \ln FL$	-0.10	-(1.84)	-0.09	-(1.64)	-0.01	-(0.07)	-0.07	-(0.90)	0.07	(2.33)	-0.13	-(1.33)
LM-1	3.06	[0.08]	3.62	[0.06]	0.46	[0.50]	1.13	[0.29]	2.82	[0.09]	2.17	[0.14]
LM-2	4.44	[0.11]	5.09	[0.08]	0.92	[0.63]	1.53	[0.46]	2.93	[0.23]	2.27	[0.32]
Normality	3.29	[0.19]	3.35	[0.18]	6.61	[0.03]	4.48	[0.10]	3.38	[1.18]	4.85	[0.08]

Notes: DTF is distance to the frontier. In the first two columns, foreign X/Q is proxied by foreign R/Y in the absence of a measure for foreign R&D workers. All growth in foreign research activity and foreign research intensity measures are weighted using import shares following the approach outlined in Eq. (A5) in the appendix. LM-1 and LM-2 are Breusch-Godfrey LM statistics for no first- and second-order serial correlation in the residuals, respectively. Normality test is based on the Jarque-Bera test which tests the null hypothesis of normal residuals. Robust standard errors are obtained using the Newey-West procedure. A first-order autoregressive term is included in all regressions. The number of observation used in the five- and ten-year differences analyses are 51 and 45, respectively. Values in parenthesis are *t*-statistics and those in square brackets are *p*-values.

The results give more support for Schumpeterian growth theory. The coefficient of domestic research intensity [$(X/Q)^d$] achieves robust economic and statistical significance when research

[3] It is worth noting that the coefficient of this variable does not show any significant economic or statistical improvement when we consider the pre-reform period of 1950-1990 (results not reported).

intensity is measured by R&D workers and real R&D expenditures, although less support for the theory is found when patent-based measures for research intensity are considered. The implied social rates of return to research intensity lie between 19 and 49 percent. These rates of return are higher than the private R&D rates of return given that domestic R&D not only promotes TFP growth, but also facilitates the assimilation of knowledge embodied in imports and thereby raise the absorptive capacity of this economy.[4] In terms of foreign research intensity [$(X/Q)^f$], our results provide quite strong support for the favorable effects of this variable on TFP growth. These results are similar to those found by Savvides and Zachariadis (2005), who use similar measures to delineate the sources of TFP growth for a panel of 32 developing economies and find convincing support for research weighted import intensity to promote TFP growth in low- and middle-income developing economies.

The coefficient of the distance to the frontier is found to be statistically and economically significant in all cases, providing convincing support to the conditional convergence hypothesis. This finding suggests that convergence has played a favorable role in lifting TFP growth in India over the long run. It is interesting to note that the coefficient on ln*DTF* lies between 0.78 and 0.92, much greater than the range of 0.02-0.10 found in Griffith et al. (2003) and 0.30-0.80 found in Madsen (2008), both of which use data for OECD economies. However, our findings are consistent with the functional assumption on convergence of Griffith et al. (2003), who postulate that any marginal increase in TFP due to conditional convergence is larger for economies that lie further away from the technology frontier. Given India's long distance to the technology frontier relative to OECD economies other than the U.S., convergence in India's TFP is expected to be much higher than OECD economies. Finally, the level of financial liberalization, which is included to partially capture the effect of economic reforms, is found to have a statistically economically significant effect on TFP growth in most cases, consistent with our prior belief. The growth rate of financial liberalization, however, is found to have little impact on TFP growth.

To shed more light on the results, we will now consider the role of absorptive capacity (CDTF), which is measured by the interaction between research intensity and distance to the frontier, in affecting TFP growth. Given the very high correlation between the research intensity measures and absorptive capacity, which pose some difficulties in making inference on the validity of each growth theory, we focus only on the results related to distance to the frontier, absorptive capacity and financial liberalization. It is evident that with few exceptions, the role of absorptive capacity is found to be economically and statistically significant. This result is in line with Griffith

[4] When considering the pre-reform period, the support for this theory is found to be relatively weaker, probably because the assumption of increasing product proliferation was not fully satisfied in India before the reform in 1991.

et al. (2004), Kneller and Stevens (2006) and Madsen (2008) who suggest that a country absorbs the technology embodied in capital goods imported from other countries faster if it devotes more of its own resources towards R&D. On the other hand, the effects of distance of the frontier (ln*DTF*) and the index of financial liberalization (ln*FL*) remain quite robust in this new set of results.

Table 6: Estimates of TFP growth with the effect of absorptive capacity

	$X = N$ $(X/Q)^d = N/L$ $(X/Q)^f = (R/Y)^f$		$X = N$ $(X/Q)^d = N/AL$ $(X/Q)^f = (R/Y)^f$		$X = R$ $X/Q = R/Y$		$X = R$ $X/Q = R/AL$		$X=PA$ $X/Q = PA/L$		$X = PG$ $X/Q = PG/L$	
5-year differences												
ln*DTF*	0.86	(9.49)	0.82	(7.77)	1.06	(6.79)	0.86	(7.74)	0.69	(6.62)	0.68	(3.40)
CDTF	1.68	(3.20)	0.11	(2.00)	-0.05	(-2.63)	-0.16	(-1.02)	0.02	(2.12)	0.07	(0.73)
ln*FL*	0.19	(3.79)	0.30	(4.94)	0.37	(3.22)	0.11	(1.31)	0.16	(3.60)	0.08	(0.59)
Δln*FL*	0.05	(1.12)	0.04	(0.81)	0.06	(0.74)	0.05	(0.94)	0.23	(4.69)	0.00	(0.02)
LM-1	0.86	[0.35]	1.29	[0.26]	1.50	[0.22]	1.02	[0.31]	2.08	[0.15]	0.37	[0.54]
LM-2	3.37	[0.19]	2.73	[0.26]	1.66	[0.44]	1.07	[0.59]	3.87	[0.14]	0.59	[0.74]
Normality	0.83	[0.65]	1.14	[0.56]	1.44	[0.48]	0.77	[0.67]	2.91	[0.23]	2.08	[0.35]
10-year differences												
ln*DTF*	0.66	(6.03)	0.66	(6.74)	1.21	(5.76)	0.98	(9.60)	0.83	(4.37)	0.41	(1.13)
CDTF	3.78	(3.32)	0.24	(4.36)	-0.06	(-2.07)	-0.52	(-1.70)	0.00	(0.07)	0.21	(1.24)
ln*FL*	-0.01	(-0.04)	0.26	(4.06)	0.62	(2.81)	-0.05	(-0.22)	0.12	(1.45)	0.19	(0.52)
Δln*FL*	-0.15	(-3.21)	-0.11	(-2.70)	-0.07	(-0.71)	-0.09	(-1.07)	0.07	(1.97)	-0.14	(-1.52)
LM-1	0.05	[0.83]	1.08	[0.30]	1.95	[0.16]	0.46	[0.50]	2.86	[0.09]	1.96	[0.16]
LM-2	0.56	[0.75]	1.56	[0.46]	1.95	[0.38]	0.86	[0.65]	2.98	[0.23]	1.96	[0.37]
Normality	5.26	[0.07]	11.17	[0.00]	7.20	[0.02]	2.12	[0.34]	3.36	[0.18]	3.03	[0.21]

Notes: results related to growth in research activity and research intensity are not reported for brevity. CDTF is the interaction term between research intensity and DTF.

6. Firm-level Analysis

The results in the previous section may have suffered from a small sample problem. Furthermore, the results using aggregate data may have been influenced by the environment under which production was undertaken in the pre-liberalization period or before 1991. Given that the Schumpeterian results are derived based on the assumption of free entry of firms, this assumption may not have been satisfied for India prior to the liberalization in 1991. Before this period, there were strict controls on entry of both domestic and foreign firms (Panagariya, 2002; Rodrik and Subramanian, 2005). These restrictions were liberalized to some extent in the late 1970s and 1980s. However, it was only after 1991 that all licensing requirements were withdrawn and the cap on foreign direct investment in most industries was lifted and raised to 51 percent. Furthermore, as Laincz and Peretto (2006) argue, Schumpeterian growth models focus on the scale of a firm and not the economy when examining steady-state growth. Thus, to complement the aggregate time series findings, we perform firm-level analysis in this section using data for the post-liberalization period.

The panel data used in this section to estimate Eq. (5) are manufacturing firm-level data for the period 1993-2005. The original data set includes a total of 3,630 firms. However, data for most of these firms are unavailable over the entire sample period. Therefore, a balanced panel of 590 firms for which data are available for the entire sample period is considered here. Note that not all of these firms invested in R&D in each year over the entire sample period. Moreover, the only measure of research activity that is available at the firm-level is R&D expenditures. Data on patents, R&D workers, and imports by sources are not available. The Appendix gives detailed accounts of the data.

6.1 Panel unit root and cointegration tests

To shed further light into the compatibility of the second-generation growth theories, we perform panel unit root and Pedroni's (2004) panel cointegration tests. The commonly used panel unit root tests are the Im *et al.* (2003) test (henceforth IPS), the ADF-Fisher type of Maddala and Wu (1999) test (henceforth MW), both of which assume an individual autoregressive unit root process, and the Breitung (2000) test that assumes a common autoregressive unit root process. However, as shown by Hlouskova and Wagner (2006), the MW test is increasingly oversized in short panels with 10 or 15 years period of data, as is the case here. Moreover, with serially correlated errors, the size distortions of the Breitung test are minimal, compared to both the MW and IPS tests. We therefore use only the Breitung's panel unit root test here. The results are reported in Table 7.

Table 7: Panel unit root tests based on the Breitung's common unit root process

Variable	$\ln A_{it}$	$\ln R_{it}$	$\ln(R/Y)_{it}$	$\ln(R/AL)_{it}$
Levels	2.912 (0.998)	-12.596*** (0.000)	-10.253*** (0.000)	-10.565*** (0.000)
First Differences	-20.901*** (0.000)	-27.094*** (0.000)	-23.498*** (0.000)	-25.220*** (0.000)

Notes: figures in the parenthesis are *p*-values. *** indicates 1% level of significance.

As with the aggregate time series data, the null hypothesis of a unit root in $\ln A_{it}$ cannot be rejected at the conventional levels of significance, whereas $\Delta\ln A_{it}$ is found to be stationary. Therefore, we can perform cointegration tests to examine whether the firm-level data are more consistent with semi-endogenous or Schumpeterian growth theory. Since $\ln A_{it}$ is $I(1)$, one would expect $\ln R_{it}$ to follow an $I(1)$ process in order to satisfy the conditions of semi-endogenous growth theory. However, our results indicate that $\ln R_{it}$ is a stationary process, and thus providing no support for semi-endogenous growth theory. On the other hand, both $\ln(R/Y)_{it}$ and $\ln(R/AL)_{it}$, are found to be $I(0)$, which is consistent with Schumpeterian growth theory.

The Pedroni's (2004) panel cointegration procedure is used to test for evidence of panel cointegration. The optimal lag length is chosen to be one in all cases based on the AIC. The results reported in Table 8 show that even though $\ln R_{it}$ and $\ln A_{it}$ appear to be cointegrated, as suggested by all seven Pedroni test statistics, the cointegrating vector does not support the predictions of semi-endogenous growth theory given that the second element in the vector is found to be non-negative.

On the other hand, there is a fairly robust support for Schumpeterian growth theory. All Pedroni's test statistics provide evidence in support of the presence of cointegration between lnR_{it} and lnAL_{it} as well as lnR_{it} and lnY_{it}. More importantly, the second elements in the cointegrating vectors are statistically significant and have the right signs, and thus providing strong evidence in support of Schumpeterian growth theory.

Table 8: Panel cointegration test

Model		Panel statistic		Group panel statistic		Cointegrating vector	
lnR_{it} and lnA_{it}	*v* statistic	3.458	(0.001)			1.000	0.037 [0.131]
	rho statistic	-15.042	(0.000)	-3.693	(0.000)		
	PP statistic	-33.618	(0.000)	-39.547	(0.000)		
	ADF statistic	-33.705	(0.000)	-37.785	(0.000)		
lnR_{it} and lnAL_{it}	*v* statistic	5.057	(0.000)			1.000	-1.563*** [-10.329]
	rho statistic	-17.111	(0.000)	-5.391	(0.000)		
	PP statistic	-35.828	(0.000)	-41.940	(0.000)		
	ADF statistic	-36.395	(0.000)	-40.811	(0.000)		
lnR_{it} and lnY_{it}	*v* statistic	3.161	(0.003)			1.000	-1.215*** [-10.700]
	rho statistic	-16.014	(0.000)	-4.258	(0.000)		
	PP statistic	-34.634	(0.000)	-41.814	(0.000)		
	ADF statistic	-35.305	(0.000)	-39.332	(0.000)		

Notes: an intercept but no trend was included in estimation. Numbers in round parenthesis are *p*-values. Figures in square brackets are *t*-statistics. *** indicates 1% levels of significance.

6.2 Panel TFP growth estimates

We estimate the following firm-level TFP growth equation that nests the predictions of both endogenous growth theories as discussed in Section 2:

$$\Delta \ln A_{it} = \varphi_0 + \varphi_1 \Delta \ln X_{it} + \varphi_2 \ln\left(\frac{X}{Q}\right)_{it} + \varphi_3 \ln\left(\frac{A^{\max_j}}{A}\right)_{it-1} + \varphi_4 \left(\frac{X}{Q}\right)_{it} \ln\left(\frac{A^{\max_j}}{A}\right)_{it-1} + TD + ID + \varepsilon_{it} \tag{6}$$

where A^{max}_j is the industry technology frontier, *TD* is a vector of time dummies and *ID* is industry dummies, which capture macroeconomic shocks including transitional shocks that affect all firms equally, and ε_{it} is independently and identically distributed errors. Distance to the frontier is measured by the difference in TFP of a firm of the j^{th} industry relative to the firm within that industry included in this sample that has the highest TFP (A^{max}_j). This is not an ideal way of measuring distance to the frontier. According to Griffith et al. (2003, 2004), one should consider the difference in TFP from the global maximum. However, data on the global maximum TFP are not readily available. Moreover, as mentioned earlier, data on sources of imports are also not available at the firm-level, which preclude estimation of the effects of international technology spillovers on TFP growth of firms.

The above equation is estimated in 5-year differences using OLS, fixed effects and system GMM dynamic panel data techniques. By taking 5-year differences, we ensure that the influence of business cycle fluctuations is filtered out. We prefer estimates based on the system GMM estimator since it sufficiently deals with endogeneity bias problems. The results reported in Table 9 suggest that none of the estimates shows a significant relationship between TFP growth and the growth in domestic R&D. The coefficient in all cases carries a negative sign, and this provides no support for semi-endogenous growth theory. On the other hand, there is some support in favor of Schumpeterian growth theory based on the preferred GMM estimator. Our results are quite consistent with those of Laincz and Peretto (2006), who also find support for Schumpeterian growth theory but not semi-endogenous growth theory based on firm-level data for the U.S.

Table 9: Panel TFP growth estimates

	OLS		Fixed effects		GMM	
	X/Q=R/Y	*X/Q=R/AL*	*X/Q=R/Y*	*X/Q=R/AL*	*X/Q=R/Y*	*X/Q=R/AL*
Intercept	-0.141	-0.158	-0.898	-0.903	0.139	-0.112
	(-4.27)	(-4.47)	(-13.69)	(-13.59)	(1.64)	(-1.78)
$\Delta \ln X_{it}$	-0.002	0.005	-0.001	0.004	-0.067	-0.005
	(0.250)	(0.790)	(0.130)	(0.490)	(1.180)	(0.270)
$\ln(X/Q)_{it}$	-0.007	-0.044	0.001	-0.064	0.522	0.051
	(1.110)	(6.260)	(0.080)	(4.920)	(1.870)	(2.560)
$\ln DTF_{it}$	0.095	0.137	0.431	0.445	0.283	0.180
	(8.010)	(8.960)	(14.280)	(14.830)	(5.080)	(3.310)
$CDTF_{it}$	0.015	0.000	0.113	0.000	2.309	0.000
	(0.470)	-(1.020)	(2.270)	(2.030)	(1.920)	(1.290)
$\Delta \ln A_{it-1}$	0.707	0.671	0.486	0.466	0.809	0.879
	(23.340)	(22.290)	(16.960)	(16.110)	(17.720)	(22.610)
m(2)					[0.571]	[0.351]
Hansen					[0.230]	[0.013]

Notes: 4130 observations were used in the OLS and fixed effects estimation, and the GMM estimation involved 2013 observations. Robust standard errors were used in all regressions. Time- and industry-dummies were included in all regressions. DTF is distance to industry frontier. $\Delta \ln A_{it-1}$ is the lagged dependant variable. m(2) is the test for no second-order serial correlation in residuals of the GMM estimation. The null hypothesis of the Hansen test is that all instruments used in the GMM estimation are valid. Different instruments have been used in the GMM regressions. Specifically, in the equation involving *X/Q=R/Y*, the set of instruments includes the lagged levels of $\Delta \ln X_{it}$ dated (*t*-2) and (*t*-3) in the first-differenced equations, and correspondingly the lagged first-differences of this variable dated (*t*-1) in the levels equation; the lagged levels of $\ln(X/Q)_{it}$ dated (*t*-3) and earlier in the first-differenced equation, and correspondingly the lagged first-differences of this variable dated (*t*-2) in the levels equation; the lagged levels of $\ln DTF_{it}$ dated (*t*-2) to (*t*-9) in the first-differenced equations, and correspondingly the lagged first-differences of this variable dated (*t*-1) in the levels equation; the lagged levels of $CDTF_{it}$ dated (*t*-4) in the first-differenced equations, and correspondingly the lagged first-differences of this variable dated (*t*-3) in the levels equation; and finally, the lagged levels of $\Delta \ln A_{it}$ dated (*t*-2) to (*t*-4) in the first-differenced equations, and correspondingly the lagged first-differences of this variable dated (*t*-1) in the levels equations. Values in round parenthesis are *t*-statistics and those in square brackets are *p*-values.

With regard to distance to the industry frontier and absorptive capacity, in most cases, the results suggest that they both promote TFP growth of a firm. This is an important result because it shows that firms that are behind the technology frontier and willing to invest in R&D will grow faster than their peers that are not investing in R&D. This implies that firms investing in R&D will experience a stronger growth in their profits than those that do not. Moreover, the joint significance of research intensity and the interaction term suggests that not only it is important for a firm to increase its R&D spending, it is also crucial for the firm to raise its research intensity in order to use the technology that is developed elsewhere effectively and to close the gap between its TFP and the technological leader's TFP. Therefore, these results have significant implications for a firm's survival and success, as also for an economy's industries. These findings indicate that although intensifying R&D growth is essential, it is not sufficient for long-term growth. What is also needed is that R&D spending keeps pace with the increasing complexity of product varieties available in the economy.

7. Conclusions and Implications of the Findings

The objective of this paper is two-fold: first, to test whether the two second-generation endogenous growth models are consistent with the data for India; and second, to test whether endogenous growth theories can explain growth in a miracle economy like India. The study is motivated by the significant increase in the GDP growth observed in India during the post-reform period, and the lack of any previous attempts in testing R&D-induced growth for developing countries.

Using aggregate time-series data over the period from 1950 to 2005 and data for 590 firms over the period 1993-2005, the results of the paper give little support for semi-endogenous growth theory. First, no robust long-run relationship between TFP and research activity is found. Second, TFP growth cannot be explained by growth in research activity. By contrast, the estimates provide strong support for Schumpeterian growth theory. In particular, there exists a long-run relationship between research activity and product lines or product varieties. The increasing number of product lines that is associated with growth in R&D activity ensures that TFP growth is not slowing down to zero or increases to infinity as predicted by first-generation endogenous growth models. History and econometric tests suggest that TFP growth is stationary in the long run. Therefore, the increasing TFP growth that India has experienced in its post-reform period does not imply that TFP growth will continue to increase. Moreover, TFP growth is positively associated with research intensity. This implies that R&D has permanent growth effects provided that R&D is continually increased to counteract the concomitant increase in the variety of products in the economy.

The estimation results also provide evidence of strong international spillover effects to the Indian economy. First, the aggregate estimates indicate significant research intensity spillovers through the channel of imports. Second, TFP growth is positively affected by India's distance to the technology frontier, enabling India to enjoy the advantage of technology backwardness. Third, India's growth is significantly affected by the interaction between research intensity and the distance to the frontier, which provides further evidence in favor of the Schumpeterian growth model of Howitt (2000). Thus, the higher the research intensity, the more opportunities for an economy to take advantage of the technology that is developed elsewhere.

The analysis points to the importance of technology as the primary source of India's TFP growth and the increasing TFP growth experienced in India during its post-reform period. The rise in research intensity that gained momentum in the post-reform period not only contributed to a higher growth rate in the Indian economy, it also facilitated the increasing use of technology developed by the technology frontier. Furthermore, the growth spurt has been facilitated by increasing transfer of research through the channel of imports. All these factors have been important for growth in the post-reform period.

These results, however, do not rule out the possibility that the post-reform growth spurt has been affected by the economic reforms that started in the 1980s. In fact, our results showed that financial liberalization has been influential for growth. Moreover, the economy has opened up as part of the reform programs and allowed a larger knowledge flow from abroad. The reforms have also given incitements to undertake R&D. For instance, weighted tax deductions and higher depreciation allowance for the purchase of R&D-related machinery and materials have been granted by the government (Government of India, 2008).

Finally, the findings in this paper have important implications for the growth prospects of India. In the long run, the Indian economy is likely to converge to the growth rates experienced in OECD economies in the post WWII period provided that it keeps it research intensity at its present level. The growth rate is likely to be higher in the medium term because India will continue to enjoy its advantage of backwardness and, therefore, be able to imitate and improve the technology that has been developed in the frontier countries. This advantage will disappear as India approaches the technology frontier, but this will take a long time.

References

Acharya, S. (2004). "India's Growth Prospects Revisited." *Economic and Political Weekly*, 39.

Aghion, P. and Howitt, P. (1992). "A Model of Growth through Creative Destruction." *Econometrica*, 60, pp. 323-351.

____ (1998). *Endogenous Growth Theory*. Cambridge: MIT Press.

Ang, J. B. (2008a). "Finance and Inequality: The Case of India." *Monash University, Department of Economics Discussion Paper Series 08/08*.

____ (2008b). *Financial Development and Economic Growth in Malaysia*. London: Routledge, forthcoming.

Ang, J. B. and McKibbin, W. J. (2007). "Financial Liberalization, Financial Sector Development and Growth: Evidence from Malaysia." *Journal of Development Economics*, 84, pp. 215-233.

Barro, R. J. and Lee, J.-W. (2001). "International Data on Educational Attainment. Updates and Implications," Centre for International Development. Harvard University.

Brahmananda, P. R. (1982). *Productivity in the Indian economy: rising inputs for falling output*. Himalaya Publishing House.

Breitung, J. (2000). "The Local Power of Some Unit Root Tests for Panel Data," Baltagi, B. (Ed), In: *Advances in Econometrics: Nonstationary Panels, Panel Cointegration, and Dynamic Panels.* Amsterdam: JAI Press, 161-178.

China-Ministry-of-Science-and-Technology (2007). "S&T Statistics." *Ministry of Science and Technology, the People's Republic of China*.

Coe, D. T. and Helpman, E. (1995). "International R&D spillovers." *European Economic Review*, 39, pp. 859-897.

De Long, B. J. (2003). "India Since Independence: An Analytic Growth Narrative," Rodrik, D. (Ed), In: *In search for prosperity: analytic narratives on economic growth.* Princeton: Princeton University Press, 184-204.

Dinopoulos, E. and Thompson, P. (2000). "Endogenous Growth in A Cross-section of Countries." *Journal of International Economics*, 51, pp. 335-362.

Gollin, D. (2002). "Getting Income Shares Right." *Journal of Political Economy*, 110, pp. 458-474.

Government-of-India (2008). "Industrial R&D Promotional Programme." *Department of Scientific and Industrial Research (2008), Ministry of Science and Technology*.

____ (various issues). "Economic Survey." *Ministry of Finance, Government of India*.

Griffith, R., Redding, S. and Reenen, J. V. (2003). "R&D and Absorptive Capacity: Theory and Empirical Evidence." *Scandinavian Journal of Economics*, 105, pp. 99-118.

____ (2004). "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries." *Review of Economics and Statistics*, 86, pp. 883-895.

Grossman, G. M. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. Cambridge: MIT Press.

Ha, J. and Howitt, P. (2007). "Accounting for Trends in Productivity and R&D: A Schumpeterian Critique of Semi-Endogenous Growth Theory." *Journal of Money Credit and Banking*, 39, pp. 733-774.

Hall, R. E. and Jones, C. I. (1999). "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, 114, pp. 83-116.

Hlouskova, J. and Wagner, M. (2006). "The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study." *Econometric Reviews*, 25, pp. 85-116.

Howitt, P. (1999). "Steady Endogenous Growth with Population and R&D Inputs Growing." *Journal of Political Economy*, 107, pp. 715-730.

____ (2000). "Endogenous Growth and Cross-Country Income Differences." *American Economic Review*, 90, pp. 829-846.

Im, K. S., Pesaran, M. H. and Shin, Y. (2003). "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics*, 115, pp. 53-74.

Jones, C. I. (1995a). "R&D Based Models of Economic Growth." *Journal of Political Economy*, 103, pp. 759-784.

____ (1995b). "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics*, 110, pp. 495-525.

Kneller, R. and Stevens, P. A. (2006). "Frontier Technology and Absorptive Capacity: Evidence from OECD Manufacturing Industries." *Oxford Bulletin of Economics and Statistics*, 68, pp. 1-21.

Korea-Ministry-of-Science-and-Technology (2007). "Statistics of R&D in Science and Technology." *Ministry of Science and Technology, Republic of Korea*.

Korea-National-Statistical-Office (various issues). "Korea Statistical Yearbook." *Korea National Statistical Office, Republic of Korea*.

Kortum, S. (1997). "Research, Patenting, and Technological Change." *Econometrica*, 65, pp. 1389-1419.

Laincz, C. and Peretto, P. (2006). "Scale Effects in Endogenous Growth Theory: An Error of Aggregation, Not Specification." *Journal of Economic Growth*, 11, pp. 263-288.

Maddala, G. S. and Wu, S. (1999). "A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test." *Oxford Bulletin of Economics and Statistics*, 61, pp. 631-652.

Madsen, J. B. (2005). "A Century of Economic Growth: The Social Returns to Investment in Equipment and Structures." *The Manchester School*, 73, pp. 101-122.

____ (2007). "Technology Spillover through Trade and TFP Convergence: 135 Years of Evidence for the OECD Countries." *Journal of International Economics*, 72, pp. 464-480.

____ (2008). "Semi-Endogenous versus Schumpeterian Growth Models: Testing the Knowledge Production Function using International Data." *Journal of Economic Growth, forthcoming*.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992). "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics*, 107, pp. 407-437.

McKinnon, R. I. (1973). *Money and Capital in Economic Development*. Washington, D.C.: Brookings Institution.

National-Bureau-of-Statistics-of-China (various issues). *China Statistical Yearbook*. Beijing: China Statistics Press.

Ng, S. and Perron, P. (2001). "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica*, 69, pp. 1519-1554.

Panagariya, A. (2002). "India's Economic Reforms: What Has Been Accomplished? What Remains to Be Done?" *ADB: ERD Policy Brief Series No. 2*.

Pedroni, P. (2004). "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis." *Econometric Theory*, 20, pp. 597-625.

Planning Commission (2007). "Macro-Aggregates," Government of India.

Rodrik, D. and Subramanian, A. (2005). "From "Hindu Growth" to Productivity Surge: The Mystery of the Indian Growth Transition." *IMF Staff Papers*, 52, pp. 193-228.

Savvides, A. and Zachariadis, M. (2005). "International technology diffusion and the growth of TFP in the manufacturing sector of developing economies." *Review of Development Economics*, 9, pp. 482-501.

Saxena, S. (2007). "Technology and Spillovers: Evidence from Indian Manufacturing Micro-Data." *Monash University, Department of Economics Discussion Paper Series 08/27*.

Segerstrom, P. S. (1998). "Endogenous Growth without Scale Effects." *American Economic Review*, 88, pp. 1290-1310.

Shaw, E. S. (1973). *Financial Deepening in Economic Development*. London: Oxford University Press.

World Intellectual Property Organization (WIPO) (2007). "WIPO Patent Report: Statistics on Worldwide Patent Activity."

Young, A. (1998). "Growth without Scale Effects." *Journal of Political Economy*, 106, pp. 41-63.

Zachariadis, M. (2003). "R&D, Innovation, and Technological Progress: A Test of the Schumpeterian Framework without Scale Effects." *Canadian Journal of Economics*, 36, pp. 566-586.

____ (2004). "R&D-Induced Growth in the OECD?" *Review of Development Economics*, 8, pp. 423-439.

Appendix: Data and Measurement Issues

Aggregate data

TFP. Based on the homogeneous Cobb-Douglas production function TFP can be recovered from the following equation:

$$TFP = \frac{Y_t}{K_t^{\alpha}(HL)_t^{1-\alpha}} \tag{A1}$$

where Y_t is real GDP, K_t is real capital stock and H_t is an index of human capital. $(HL)_t$ measures the quality adjusted workforce and α measures capital elasticity. The assumption of constant returns to scale is maintained. K_t is constructed using the standard perpetual inventory method as follows:

$$K_{t+1} = K_t(1-\delta) + I_{t+1} \tag{A2}$$

where δ is the depreciation rate and I_t is investment. Two different types of capital are considered: non-residential structures (construction) and machinery. They are deflated using the gross fixed capital formation deflator to express in real terms. Following standard practice in the literature (see, e.g., Coe and Helpman, 1995), the initial capital stock (K_0) for each type of capital is determined as follows:

$$K_0 = \frac{I_0}{\delta + g} \tag{A3}$$

where I_0 is the investment in physical capital in the initial period under consideration and g is the average geometric growth rate over the period 1950-2005. The rate of depreciation (δ) is assumed to be 3 percent for non-residential structures and 10 percent for machinery. In the literature, the standard depreciation rate for construction is 3 percent, and for machinery for developed countries is 17 percent (see Madsen, 2005).

While these data are obtained from the National Accounts Statistics (NAS), labor force (L_t) data are not available from domestic sources. Hence, they are compiled from the Penn World Tables. Unfortunately, data on hours-worked in India are unavailable and therefore they are not corrected for hours-worked. However, piecemeal data on hours-worked from the International Labor Organization (ILO) suggest that they have varied from 43 to 48 hours per week across different economic activities with little variation over time. Therefore, this is unlikely to bias our TFP estimate significantly.

Human capital (H_t) is measured by assuming a piece-wise linear rate of return of 13.4 percent for the first four years of schooling and 10.1 percent for the subsequent four years (see Hall and Jones, 1999). In the equation below, r is the average return to schooling and s is the average years of schooling for population aged 25 and above:

$$H = (1+r)^s \tag{A4}$$

The data are gathered from Barro and Lee (2001), and the missing data are interpolated and extrapolated using an exponential growth rate. The inclusion of human capital refines TFP estimates since it has been found to be a significant contributor to growth (Hall and Jones, 1999). Moreover, this produces a more reliable estimate of physical capital that is closer to theoretically acceptable levels of 0.3-0.35, as found by Mankiw *et al.* (1992).

There are two methods that can be used to determine labor's income share in total output. The first is to use the sum of *compensation to employees* and *mixed income* available from NAS and divide by total output. *Mixed income* includes incomes of the self-employed and rent/profit accruing to unorganized/informal enterprises, and therefore, have to be counted as labor's income rather than profits, as argued by Gollin (2002). These data are available from 1980 onwards. Earlier estimates of factor income could be obtained from Brahmananda (1982) who has estimated this share to be 75 percent between 1950 and 1970, and 71 percent between 1970 and 1980. The other method is to assume a constant share of 0.7

following the conventional practice. We have considered both methods in the estimation, and they give qualitatively similar results. For simplicity we have chosen the second method.

Research activity. Research activity in this study is measured by R&D personnel, R&D expenditures, patents granted to domestic residents and patent applications by domestic residents. Data on the R&D-based measures are gathered from various publications on "R&D Statistics" of the Department of Science and Technology, Government of India and Planning Commission (2007). These data are complemented with various issues of the *Statistical Yearbook* published by the United Nations Educational, Scientific, and Cultural Organization (UNESCO). Data on R&D expenditures are available at five year intervals between 1950 and 1970, and continuously thereafter. Missing data are interpolated using a geometric growth rate.

Following Madsen (2008), nominal R&D expenditures are deflated using two different deflators. The first deflator is constructed an unweighted average of the labor costs deflator and GDP deflator. The second deflator takes into consideration that besides labor inputs, R&D also requires materials, equipment and structures. Therefore, the second R&D deflator is constructed using labor costs deflator (45%), GDP deflator (45%), equipment (5%) and structures (5%) deflator respectively, where the values in brackets represent weights. The estimated results pertain only to real R&D data based on the first deflator although the results remain robust to use of either deflator. Note that the data on total R&D personnel and R&D expenditures also include defense-related R&D. A distinction between defense-related and civilian R&D personnel and expenditures cannot be made due to data limitations.

Data for R&D personnel between 1950 and 1990 are available at ten year intervals, and continuously thereafter. We use the stock of engineers (obtained from the University Grants Commission of India) to generate sufficient variations in the series while interpolating the missing years. Patent data are obtained from the World Intellectual Property Organization (2007). One of the foremost advantages of using patent data as measures of R&D activity is that they directly measure research output, and therefore do not require any normalization when considering Schumpeterian growth theory. Nonetheless, the disadvantage of using patent data is that not all innovations are patented. Moreover, the average value of patents may have changed over time.

International technology spillovers are measured by an import-ratio weighting scheme as follows:

$$\left(\frac{X}{Q}\right)^f_{it} = \sum_n \frac{m_{ijt}}{m_{it}} \left(\frac{X}{Q}\right)^d_{jt} , \qquad i \neq j$$

$$X^f_{it} = \sum_n \frac{m_{ijt}}{m_{it}} \bar{X}^d_{jt} , \qquad i \neq j \qquad \text{(A5)}$$

where d and f stand for domestic and foreign respectively, n is the number of India's import partners, m_{ij} is India's import of high-technology products from country j, m_i is India's total imports of high-technology products, $\bar{X}$ is an index of innovative activity, which takes on a value of one in 2000. The following SITC classifications are used for high-technology products: chemicals and related products (SITC Section 5), machinery and transport equipment (SITC Section 7), and professional and scientific instruments (SITC Section 8.7). All countries which have had a larger than 0.5 percent share in India's total imports are included in these estimates. Data on India's trading partners are gathered from various publications of the International Monetary Fund and the United Nations. Albeit it is found that the Russian Federation (former Soviet Union) held a major share in India's imports over the period 1965 and 1985, it is not included in the construction of international technology spillovers as most of India's imports from Russia are defense related equipment, which contribute little towards TFP growth. Data on TFP level, employment and research-intensity of all of India's trading partners are obtained from Madsen (2008) except for China and South Korea. Data for these two countries are collected from National Bureau of Statistics of China (various issues), Korea National Statistical Office (various issues), China Ministry of Science and Technology (various issues) and Korea Ministry of Science and Technology (various issues). Data on imports by country and product category are taken from UN Comtrade Database (2007). However, this information is only available from 1962 onwards. For prior years, the trade-value has been assumed to be the same as in 1962. Data on SITC 8.7 is available from 1978 onwards. For prior years, the value has been assumed to be the same as in 1978. Both the weighted foreign R&D stock variables have been multiplied by India's import intensity, mi. Data on India's imports are taken from Government of India (various issues)

Financial liberalization index. This measure is estimated using the principal component method by Ang (2008a). The approach considers nine indicators of financial repressionist policies. Six of them are interest rate controls, including a fixed lending dummy, a minimum lending rate, a maximum lending rate, a fixed deposit dummy, a minimum deposit rate and a maximum deposit rate. These policy controls are translated into dummy variables which take the value of 1 if a control is present and 0 otherwise. The remaining three policies are directed credit programs, the cash reserve ratio and the statutory liquidity ratio. The extent of directed credit programs is measured by the share of directed credit lending in total lending. The other two variables are direct measures expressed in percentages. These policy variables are summarized into an overall measure of financial liberalization using the method of principal component analysis. The inverse of this measure can be interpreted as the extent of financial liberalization (see, e.g., Ang and McKibbin, 2007; Ang, 2008b).

Firm-level data

The panel data used for this research are manufacturing firm-level data for the period 1993- 2005. This dataset is taken from *Prowess*, an electronic database maintained by the Centre for Monitoring Indian Economy in Mumbai. *Prowess* compiles information from the annual reports of large- and medium-sized Indian firms including government undertakings, whose shares are regularly traded on major Indian stock exchanges. It covers nearly 5,000 manufacturing firms over the entire thirteen year period, encompassing all manufacturing industries. These firms collectively account for 70 percent of total value added by the Indian manufacturing industry. However, *Prowess* does not account for very small firms from the informal/unregistered sector. Therefore, any industry which is dominated by small scale firms is not adequately represented in this dataset. Notwithstanding this limitation, there is considerable variation in the size of firms for all industries included in the dataset. We measure the TFP for each firm as follows:

$$TFP_{it} = \frac{Y_{it}}{K_{it}^{1-\alpha_1-\alpha_2} L_{it}^{\alpha_1} M_{it}^{\alpha_2}} \qquad \text{(A6)}$$

where Y_{it} is real gross output for firm i at time t, K_{it} is the capital stock, L_{it} is the number of workers and M_{it} is real material consumption. α_1 and α_2 are the factor-shares of labor and materials, respectively. Factor shares are calculated as the ratio of factor incomes to total output at current prices. Deflated measures of output and inputs have been derived using the methodology adopted by (Saxena, 2007).