MONASH University
Business and Economics

Department of Economics

Issn 1441-5429

Discussion paper 03/09

THE INDIAN GROWTH MIRACLE AND ENDOGENOUS GROWTH

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

Abstract:

Using over half a century of R&D data for India, this paper tests whether the second-generation endogenous growth theories are consistent with India's growth experience. Furthermore, the paper also examines the extent to which growth in India can be explained by R&D activity, international R&D spillovers, catch-up to the technology frontier and policy reforms. The empirical results show that the growth in India over the past five decades has been significantly driven by research intensity following the predictions of Schumpeterian growth theory.

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: Helpful comments and suggestions from workshop participants at Monash University, conference participants at 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 and, particularly, two referees are gratefully acknowledged. James B. Ang and Jakob B. Madsen acknowledge support from ARC Discovery Grants from the Australian Research Council.

The Indian Growth Miracle and Endogenous Growth

1. Introduction

Although R&D has been assumed to be the key factor 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 a "miracle" economy. 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 remains largely unexplored. In fact, the implications of R&D-based theories for developing economies remain unclear. Using data for India for over half a century, the main objectives of this paper are: 1) to test which second-generation endogenous growth theory is most applicable in explaining growth in India; and 2) to examine the importance of R&D, among other variables, in explaining growth rates in India. The growth theories tested in this paper are the following two second-generation 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), Peretto (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 endogenous growth models for a developing country. An important question addressed in this paper is whether R&D activities play a crucial role in explaining growth for a developing country like India or whether R&D-driven growth is limited to highly developed countries. The analysis is limited to India because 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 different second-generation growth models. The R&D data for India cover the period from 1950 to 2005, a time span that even exceeds that of the R&D data that are available for almost all OECD countries.[1] Most R&D data for developing countries are available only from the 1980s, which is far too short a period to discriminate between growth models using aggregate data.

Furthermore, India is an ideal candidate for testing R&D driven growth given that it has experienced a significant increase in the growth rate of total factor productivity (henceforth, TFP) since the 1980s. This raises the question of whether R&D or economic reforms have been the key factors behind India's take-off in the late 1980s. There is a 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. Hence, the literature often argues that the high growth rates have been driven by

---

[1] 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).

economic liberalization or an attitudinal change favoring the reforms (Panagariya, 2002; De Long, 2003; Rodrik and Subramanian, 2005).

Another possibility is that growth in India has been predominantly driven by transitional dynamics and factor accumulation over the past two decades. Young (1995), for example, finds that capital deepening, increasing educational attainment, and increasing labor force participation rates explain most of the high growth rates experienced among the Four Tigers in the post WWII period. However, factor accumulation is unlikely to have been entirely responsible for the high growth rates recently experienced in India for two reasons. First, the growth in India's savings rates has not coincided with the increasing productivity growth rates, as experienced by the Four Tigers. In fact, the growth in India's saving rate has fluctuated around a constant level over the period 1950-2005. If savings-induced capital deepening was the principal factor behind productivity growth in India, we would expect growth to have been distributed evenly over the period 1950-2005 and not concentrated on the latter part of the period. Second, the factor accumulation hypothesis predicts that the growth in the capital-output ratio precedes growth in labor productivity. However, Granger causality tests suggest that the capital deepening has been a result of productivity growth and not the other way around.[2] Finally, Madsen and Ang (2009) show that more than half of per capita growth in India during the period 1950-2006 is explained by TFP and that most growth in the post-reform period is driven by TFP growth.

To account for the influence on TFP growth of the liberalization attempts by the Indian Government, we include several control variables that capture the impact of the economic reforms, including an index of financial liberalization, tariff rates, the fraction of firms that are foreign owned and patent rights protection. This will also shed some light on the underlying factors behind TFP growth in India, an issue that remains largely unexamined. Both aggregate data (1950-2005) and firm-level data (590 firms over the period 1993-2005) are used in this paper to provide insight into the ability of the second-generation growth models in explaining the productivity growth experience for India.

The next section contains a brief anatomy of the most important reforms undertaken in India since independence with special focus on R&D policies. The innovation-driven endogenous growth theories and their growth implications are presented and briefly discussed in Section 3. Data and graphical evidence are presented in Section 4. Time series tests using aggregate data are performed in Section 5. Section 6 provides estimates of TFP growth. In addition to discriminating between the

---

[2] Ganger causality test of the null hypothesis that the growth in the *K-Y* ratio does not Granger cause labor productivity growth yields an *F*-statistic of 1.45 (p = 0.23). The null hypothesis that labor productivity growth does *not* precede the *K-Y* ratio is strongly rejected (*F*-statistic = 7.26, p = 0.009). These results suggest that capital deepening has been the result and not the cause of productivity growth. The Granger causality tests are estimated with one lag, which was chosen based on the likelihood ratio tests, using annual data over the period 1950-2005.

second-generation endogenous growth models, the empirical analysis also examines the roles of international technology spillovers and distance to the frontier in driving productivity growth. Some robustness checks of the estimates are provided in Section 7. Analysis using firm-level data are undertaken in Section 8 to complement the results based on aggregate data. The last section concludes the paper.

2. From Hindu to Miracle Growth Rates

Over the period from 1950 to 1990, India's per capita income grew at an average annual rate of only about 2%. During this period, the Indian government implemented restrictive trade, financial and industrial policies. Shortly after independence the Indian state took control of major heavy industries and private firms were only allowed to enter a few consumer and intermediate industries while being subject to widespread industrial licensing requirements and price regulation (Aghion et al., 2008). Furthermore, exit of firms was constrained and the liquidation of assets was rendered difficult. Additional controls on large companies were introduced in the 1960s. These include additional licensing requirements, capacity restrictions and restrictions on types of products they could produce. Trade policies were oriented towards self-sufficiency, following the foreign exchange crisis in the mid 1950s, by the introduction of the Import Trade Control Act of 1955. Following this act, imports of consumer goods were prohibited and heavy tariffs were introduced on almost all goods. In terms of financial sector policy, from the beginning of the 1960s the government gradually tightened control over the financial system by raising statutory liquidity and cash reserve requirements that continued throughout the 1970s and 1980s.

Interestingly, the Indian government promoted R&D by fostering commercialization of research undertaken by governmental funded bodies after Independence. Furthermore, several training schools for engineers and scientists were established. However, the incentives to produce in-house research were probably weak since there was little or no threat to incumbent firms from entry of domestic or foreign firms or from imported products. Furthermore, there was little expansion of product variety due to the enforcement of licensing requirements. International technology transfer was also limited since foreign partnership and technology imports were strictly regulated, especially during the 1960s (Joshi and Little, 1996). On balance, compared to the standard of other developing countries, the research intensity of Indian companies during that period was not particularly low. The average R&D expenditure as a percentage of GDP was 0.4% over the period 1950-1990.

The restrictions on trade and production were gradually lifted and capital markets were liberalized in a series of reforms that started in the late 1970s and gained strong momentum in the early 1990s. The objective was to restructure the entire orientation of India’s development strategy and to give more scope for markets in price determination and resource allocation. Consequently, the interest rates were gradually liberalized, reserve and liquidity ratios were reduced significantly, industrial licensing requirements were abolished for most industries, R&D was strongly promoted, and several trade restrictions were removed, including marked reductions in tariff rates and import quotas.

Alongside these developments, several incentive schemes have been launched to stimulate in-house R&D. Fiscal incentives to undertake R&D were already introduced in 1973 and import duties on capital goods used for R&D were also reduced. Policies to promote R&D were significantly strengthened in 1990s. The patent protection framework was strengthened substantially and several tax concessions were given to firms when undertaking in-house R&D. For instance, in 1997, a weighted tax deduction of 125 percent was allowed on purchase of capital equipment for R&D purposes. Today, any in-house R&D expenses enjoy weighted tax deductions of 150 percent. Furthermore, accelerated depreciation for R&D equipment, generous exercise duty waivers, tax holidays for R&D intensive firms, and tax incentives for R&D collaborations between private firms and universities have also been introduced. The period 1991-2005 saw a significant increase in R&D activity with an average R&D in percentage of GDP of 0.8%, two times larger than that of the pre-reform period.

The economy grew at an average annual growth rate that exceeded 6% in per capita terms in the period 1990-2005, which is 4 percentage points higher from the Hindu growth rate experienced in the period 1950-1990. The coincidence of increasing growth rates and reforms in the 1990s has led a large body of the literature to argue that the policy reforms have been the main drivers of the increasing growth rates. However, very little work has been done in quantifying the policy shifts and their effects on growth, and even less attention has been given to the influence on growth of R&D (Acharya, 2004). The next sections seek to uncover the effects of R&D and reform policies on growth, and test whether the effects of these policies and R&D on growth are temporary or permanent. As shown in the next section, the shape of the ideas production function determines whether R&D has permanent or temporary growth effects.

3. R&D-Based Endogenous Growth Models

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

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

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

where g_A is TFP growth, A is the levels of TFP, X represents research input (semi-endogenous growth theory) 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 endogenous growth models lies in the values of ϕ and β. Semi-endogenous growth theory predicts $\phi < 1$ and $\beta = 0$ whereas Schumpeterian growth theory posits $\phi = 1$ and $\beta = 1$, and the 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 rates in these economies have not increased during 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 constant returns to scale in ideas production, and replace it with the assumption of diminishing returns to knowledge at the aggregate level of the economy. 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), Peretto (1998), Howitt (1999) and Dinopoulos and Thompson (2000) retain the assumption of constant returns to knowledge from the first-generation 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 activity. 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. 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. As Laincz and Peretto (2006) point out, Schumpeterian growth models eliminate the economy-wide scale effects because they focus on the scale of the firm, and not the scale of the economy.

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, 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 $\Delta \ln A_t$ is not stationary, as predicted by the first-generation endogenous growth models for economies with increasing R&D, 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 discussed later in Section 5, the second-generation endogenous growth theories imply that the terms υ_t and ς_t in the following equations are stationary:

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

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

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) that both $\ln X_t$ and $\ln A_t$ to be non-stationary and integrated at the same order; and ii) that both variables 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., $\phi < 1$. However, if the data are consistent with Schumpeterian growth theory, one would expect: i) $\ln(X/Q)_t$ to be stationary so that measures of R&D intensity exhibit no large persistent movements; and ii) that both $\ln X_t$ and $\ln Q_t$ to be cointegrated with a cointegrated vector of $(1, -1)$ 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 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 predictions of second-generation endogenous growth models. For example, if 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):

$$\Delta \ln A_t = \gamma_0 + \gamma_1 \Delta \ln X_t^d + \gamma_2 \left(\Delta mi_t \ln X_t^f\right) + \gamma_3 \ln\left(\frac{X}{Q}\right)_t^d + \gamma_4 mi_t \ln\left(\frac{X}{Q}\right)_t^f$$
$$+ \gamma_5 \ln\left(\frac{A^{US}}{A^{IN}}\right)_{t-1} + \gamma_6 \ln FL_t + \gamma_7 \Delta \ln FL_t + e_t \quad (5)$$

where the superscripts d and f stand for domestic and foreign, mi_t is import penetration defined as the ratio of imports to GDP, $(A^{US} / A^{IN})_{t-1}$ measures the distance to the technology frontier, A_t^{US} and A_t^{In} are the TFP levels of the US and India, FL_t is an index of financial liberalization, and e_t is a stochastic error term. While semi-endogenous growth theory predicts γ_1, $\gamma_2 > 0$ and γ_3, γ_4, $\gamma_5 = 0$, Schumpeterian growth theory predicts γ_3, γ_4, $\gamma_5 > 0$ and γ_1, $\gamma_2 = 0$. It is expected that γ_6 and $\gamma_7 > 0$. Thus, R&D have only temporary effects on TFP growth under semi-endogenous growth theory while R&D has permanently positive growth effects in the Schumpeterian growth models as long as R&D is kept in a constant proportion of the number of product lines or product varieties.

The model allows international technology spillovers through the channel of imports due to Coe and Helpman (1995) and Madsen (2007b, 2008). Here imports of technology is multiplied by import penetration so that the variables ΔX^f and $(X/Q)^f$ are not influenced by the propensity to import. While ΔX^f and $(X/Q)^f$ are constructed using import weights the weights add up to one as shown below. Following Coe and Helpman (1995) it is plausible that technology spillovers through the channel of imports are proportional to the propensity to import. In other words, the larger is the propensity to import, the higher is the opportunity of domestic producers to take advantage of the technology that is produced elsewhere. 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 R&D and the R&D embodied in imported intermediate inputs. Hence, technology is transmitted internationally by the import-weighted stock of knowledge.

The above specification allows for the possibility that growth is enhanced by the distance to the technology frontier (Howitt, 2000; Griffith et al., 2003). The larger is the technology gap the more a country can take advantage of the technology developed by the frontier country because the effective costs associated with the development of a new product declines with technological distance. Finally, an index of financial liberalization is included in the model to control for the potential positive growth effects of financial liberalization. Financial liberalization is expected to affect the level or the growth in productivity positively because it is associated with greater mobilization of savings and more efficient allocation of resources (McKinnon, 1973; Shaw, 1973). The financial liberalization index is entered in levels as well as in differences to allow for the possibility that financial liberalization has temporary and permanent growth effects.

4. Data

The models are estimated using aggregate data and firm-level data. Annual data for the period 1950-2005 are used in the aggregate analysis and the firm-level analysis considers the post-liberalization period of 1993-2005. TFP_t is computed as $Y_t/[K_t^{\alpha} \cdot (HL)_t^{1-\alpha}]$ where Y_t is real GDP, K_t is real capital stock, L_t is labor force and H_t is an index of human capital. $(HL)_t$ measures the quality adjusted workforce and α measures the capital elasticity. The assumption of constant returns to scale in production is maintained and capital's share of income (α) is set at 0.3. The estimates of K_t are based on the perpetual inventory method where the initial capital stock for each type of investment (non-residential structures and machinery) is estimated using the gross capital formation in 1950 divided by the depreciation rates (3% for non-residential structures and 10% for machinery) and the average growth in gross capital formation over the period 1950-2005. National account statistics are obtained from Government of India (various issues-a) whereas labor force data are derived from the Penn World Table.

Human capital (H_t) is based on piece-wise linear rate of return to schooling 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). It is measured as $H_t = (1+r)_t^s$ where r is the average return to schooling and s is the average years of schooling for the population aged 25 and above, which is obtained from Barro and Lee (2001).

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 tests 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 $(PG/L)_t$. Here $(N/AL)_t$ and $(R/AL)_t$, are adjusted by TFP to allow for the increasing complexity of new innovations as the economy advances (Aghion and Howitt, 1992). Real output

(Y_t) and the labor force (L_t) are used as measures of product variety following Zachariadis (2003), Griffith et al. (2004), Ha and Howitt (2007) and Madsen (2008). The labor force is often used as a measure of product variety because the number of products is equal to the size of the population in steady state in Schumpeterian growth models.[3]

Following Coe and Helpman (1995) and Madsen (2008), nominal R&D expenditures are deflated using the unweighted average of the labor costs deflator and the GDP deflator.[4] Data on the R&D-based measures are gathered from various publications on "R&D Statistics" of the Department of Science and Technology (see Government of India, various issues-b) 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). Patent data are obtained from the World Intellectual Property Organization (2007).

International technology spillovers are measured by an import-ratio weighting scheme adapted from Coe and Helpman (1995), given as follows:

$$\left(\frac{X}{Q}\right)_{It}^{f} = \sum_{n} \frac{m_{Ijt}}{M_{It}} \left(\frac{X}{Q}\right)_{jt}^{d}, \quad I \neq j$$
$$X_{It}^{f} = \sum_{n} \frac{m_{Ijt}}{M_{It}} \bar{X}_{jt}^{d}, \quad I \neq j \tag{6}$$

where n is the number of import partners for India, m_{Ij} is India's import of high-technology products from country j, M_I is India's total imports of high-technology products, $\bar{X}_j$ is the real R&D for country j, which takes on a value of one in 2000.[5] Data on total imports and imports by product category from these countries are collected from various publications on international trade of the IMF, United Nations and the UN Comtrade database. Data on TFP levels, employment, R&D and research-intensity of all 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

[3] Another direct measure of product variety is import or export variety data (see Feenstra, 1994; Feenstra et al., 1999; Chen and Feenstra, 2008; Feenstra and Kee, 2008; Frensch and Wittich, 2009). In Feenstra et al. (1999), product variety for Korea and Taiwan is measured using the disaggregate exports data from these countries to the US. This approach is not considered here since the export sector in India is relatively small compared to Korea and Taiwan. Its average share of exports in GDP from the period 1975 to 2005 was only 7.66%, compared to 28.60% in Korea and 50.51% in Taiwan. Thus, unlike the cases of Korea and Taiwan, export variety is unlikely to be a good proxy for economy-wide product variety for India. Moreover, data on product variety do not go sufficiently far back in time, particularly not for our measure of spillovers of research intensity through the channel of imports. The data are available from 1975 for India and later for some of India's trade partners.

[4] 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.

[5] 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 that have a larger than 0.5 percent share in India's total imports are included in the analysis.

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).

The index of financial liberalization is obtained from Ang (2009). 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 that 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 cash reserve ratio and the statutory liquidity ratio are direct measures of financial repression and are expressed in percentages. These policy variables are summarized into an overall measure of financial liberalization using principal component analysis. The inverse of this measure can be interpreted as the extent of financial liberalization (see Ang and McKibbin, 2007).[6]

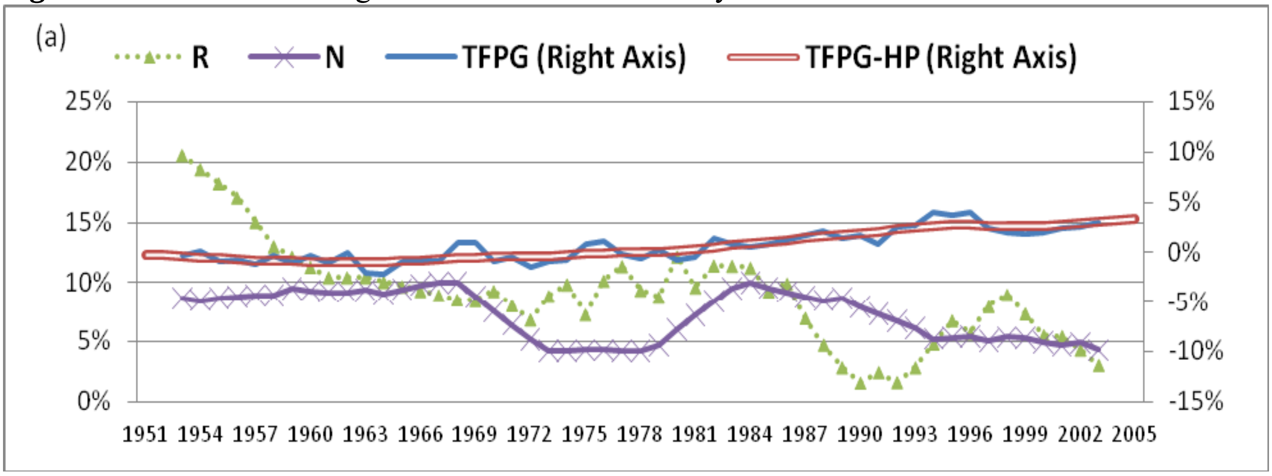
The panel data used in this section to estimate the TFP growth equation are annual manufacturing firm-level data covering the period 1993-2005. This dataset is from *Prowess*, an electronic database maintained by the Centre for Monitoring Indian Economy in Mumbai. *Prowess* compiles information from the annual reports of large- and medium-sized Indian firms including government undertakings, whose shares are regularly traded on the major Indian stock exchanges. These firms account for 70 percent of total value added by the Indian manufacturing industry. 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 in the analysis. Note that not all of these firms invested in R&D in each year over the entire sample period, and the only measure of research activity that is available at the firm level is R&D expenditures. TFP for the firm-level analysis is estimated by $Y_{it} /(K_{it}^{1-\alpha1-\alpha2} L_{it}^{\alpha1} M_{it}^{\alpha2})$, where M_{it} is real material consumption. $\alpha1$ and $\alpha2$ 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 (see Saxena, 2009 for more details).

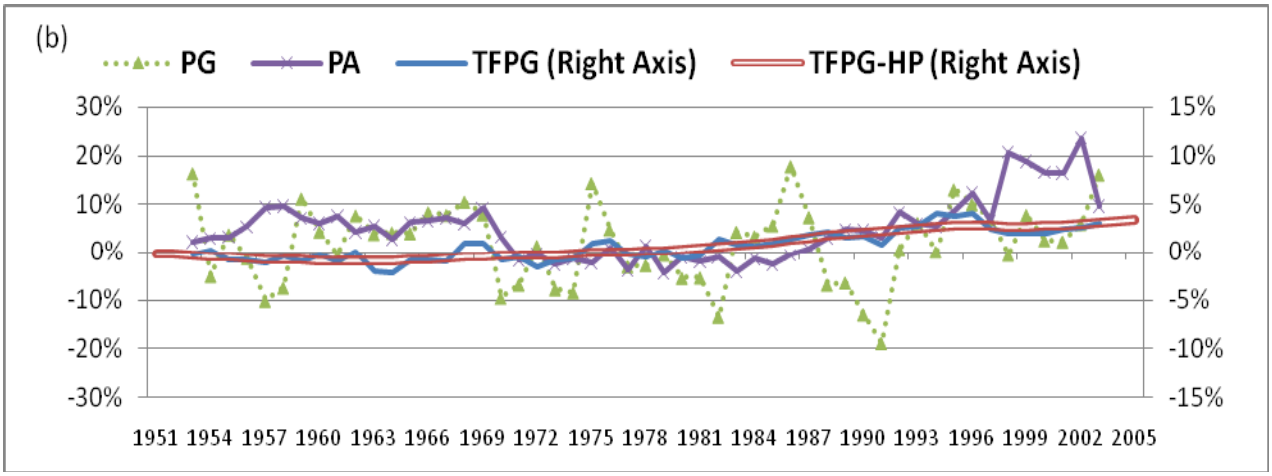
Figures 1a and 1b compare TFP growth with the growth rates in R&D workers employed, real R&D expenditures, patent applications and patents granted. If the prediction of semi-endogenous growth theory is valid, one would expect TFP growth to move closely with the growth rates of R&D activity over time. The trend in the TFP growth rate has been increasing since the 1970s. However, except for patent applications, the growth rates the R&D activity measures do not display an increasing trend. The growth rates of R&D workers and real R&D expenditures show a

[6] For more details on construction of variables and data sources, see Appendix V.

downward rather than an upward trend (Figure 1a). Furthermore, correlation between the growth rates in TFP and patent applications is small or absent. Overall, the graphical evidence presented here gives only limited support to semi-endogenous growth theory.

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

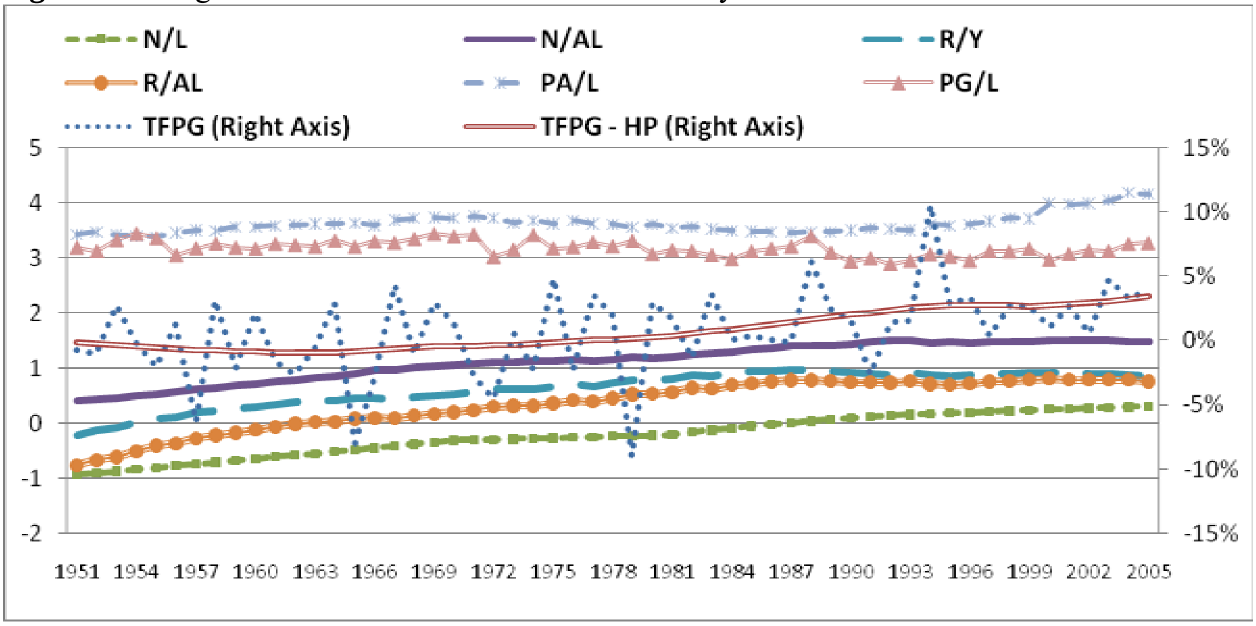




Notes: TFPG is TFP growth rate, *R* is real R&D expenditures; *N* is number of scientists and engineers engaged in R&D; *PA* is patent applications from residents and *PG* is patents granted to residents. Growth in TFP and all measures of research activity are based on 5-year centered moving average. TFPG-HP is the Hodrick-Prescott filtered series for TFP growth with a smoothing parameter of 100.

Figure 2 shows 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 an increasing TFP growth rate. Except *N*/*L* which has 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* appears to move pro-cyclically with TFP growth. Although the ratio of *PA*/*L* does not seem to exhibit any strong correlation with TFP growth, it has continued to rise after the reform period, reflecting the increasing drive towards innovation during this period. Overall, the graphical evidence gives some support to Schumpeterian growth theory.

Figure 2: TFP growth and measures of research intensity



Notes: R 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, and 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.

5. Integration and Cointegration Analyses

5.1 Unit root tests

Unit root tests are performed based on the procedure of Ng and Perron (2001) that takes into account the possible presence of a structural break using data over the period 1950-2005.[7] The details of the unit root results are reported in Madsen et al. (2009). The results are summarized here as follows. The null hypothesis of a non-stationary TFP level is not rejected but that of TFP growth is rejected at the one percent significance level. Since TFP levels follow an $I(1)$ process, one would expect R&D activity measures (i.e., $\ln N_t$, $\ln R_t$, $\ln PA_t$ and $\ln PG_t$) to contain a unit root as well to satisfy the predictions of semi-endogenous growth theory of cointegration between ln$A$ and ln$X$. However, unit root tests show no evidence that the R&D inputs follow an $I(1)$ process. This suggests that TFP has not been driven by a proportional increase in research activity as predicted by semi-endogenous growth theory. Regarding research intensity, and 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 $(PG/L)_t$ suggest that these R&D intensity measures are stationary except $\ln(N/L)_t$. Coupled with the finding of the stationarity of TFP growth, these results are consistent with the predictions of Schumpeterian growth models.

---

[7] 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 ests 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).

5.2 Cointegration tests

5.2.1 Semi-endogenous growth theory

The Johansen approach is used to examine the existence of a long-run relationship between TFP and R&D activity as predicted by semi-endogenous growth theory (Eq. 3). 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 1 presents the results of the Johansen cointegration tests and the cointegrating vector associated with each bivariate VECM. While cointegration is found only in two cases for the pre-reform period, the results of both the trace and maximum eigenvalue tests unanimously suggest that a cointegrated relationship exists in all cases for the whole sample period. These results appear to be inconsistent with the unit root tests obtained earlier.

However, 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 the estimated coefficient of $\ln A_t$ in the remaining three cases has the correct negative sign, it is only statistically significant in the third model for the full sample period. Furthermore, the magnitude of the coefficient appears unreasonably large (-74.648). Thus, the results cast doubts on the assumption of diminishing returns to knowledge as maintained by semi-endogenous growth models.

Table 1: Johansen cointegration tests for semi-endogenous growth theory (Eq. (3)).

	Pre-reform (1950-1990)			Full sample (1950-2005)		
Model	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 the variables are not cointegrated. The estimation includes an intercept but not a deterministic trend. Figures in square brackets are *t*-statistics. *, ** and *** indicate 10%, 5% and 1% levels of significance, respectively.

5.2.2. Schumpeterian growth theory

The Johansen approach is again used to examine the existence of a long-run relationship between R&D activity and product variety (Eq. 4) over the pre-reform period 1950-1990 and over the full sample period 1950-2005. The results reported in Table 2 show that the null of no cointegration is rejected at conventional significance levels in almost all cases. The second elements in the cointegrating vectors are statistically significant and have the signs as predicted by the theory in eight of the 12 cases. Furthermore, when trademarks are used as measures of product variety all cointegration estimates give support for Schumpeterian growth theory (see Appendix III). Overall, one can conclude from these cointegration tests that the Indian aggregate data are consistent with the predictions of Schumpeterian growth theory.

Table 2: Johansen cointegration tests for Schumpeterian growth theory (Eq. (4)).

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 $\ln L_t$	13.481* (13.039*)	1.000	-3.671*** [-24.075]	20.036** (19.177***)	1.000	-3.550*** [-35.942]
$\ln N_t$ and $\ln A_tL_t$	6.224 (6.021)	1.000	-2.031** [-1.774]	15.494** (14.264**)	1.000	1.787 [1.561]
$\ln R_t$ and $\ln A_tL_t$	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: see notes to the previous table.

However, the evidence gathered so far is not adequate to provide full support for Schumpeterian growth theory. Note that none of the estimated coefficients of product variety measures are close to minus one as predicted by Schumpeterian growth theory. Although these results are consistent with the literature (see, e.g., Ha and Howitt, 2007; Madsen, 2008), they are likely to reflect that product variety cannot be measured precisely. Moreover, although unit root tests suggest that TFP growth rate follows a stationary process in our sample period from 1950-2005, standard growth models predict a much longer transitional period of at least 70 years is required to achieve steady-state equilibrium. The relatively short sample period used in our analyses therefore may not provide credible inferences regarding the steady-state behavior predicted by Schumpeterian growth theory. The lack of a precise measure of product variety and the issues of steady-state behavior suggest that we should interpret the results with some caveats in mind. Some of these issues will be addressed in Sections 6 and 7.

6. TFP Growth Estimates

The TFP growth equation given by Eq. (5) is estimated in 5- and 10-year differences to filter out the influence of business cycles. The research intensity measures are taken to be the average over all years covered by the differences. Distance to the frontier is evaluated at the first year of the differences. The standard errors are derived based on the Newey-West procedure in order to provide heteroscedasticity and autocorrelation consistent estimates.

The estimation results are reported in Table 3. First, consider the estimates related to semi-endogenous growth theory. The estimated coefficient of growth in domestic research activity ($\Delta \ln X^d$) is significantly negative in more than half of the regressions. Only one of the 12 cases is the coefficient of ($\Delta \ln X^d$) positive and significant. These results provide strong evidence against semi-endogenous growth theory.[8] Furthermore, the estimated coefficients of the growth in knowledge spillovers through the channel of imports ($\Delta mi \ln X^f$) are insignificant in all cases, which provide further evidence against semi-endogenous growth theory.

The results give more support for Schumpeterian growth theory. The estimated coefficients of domestic research intensity [$(X/Q)^d$] are economically and statistically highly significant when research intensity is measured by R&D workers and real R&D expenditures in most cases, although less support for the theory is found when patent-based measures for research intensity are considered. Support for Schumpeterian theory is also found when product varieties are measured by trademarks (see Appendix III). The estimated coefficients of $mi \ln(X/Q)^f$ is statistically and economically highly significant in approximately half of the cases, which gives further support for Schumpeterian theory. The latter result is in line with the findings of Savvides and Zachariadis (2005), who use imports of research intensity from the G5 countries for a panel of 32 developing economies to explain TFP growth. They find that import intensity promote TFP growth in low- and middle-income developing economies.

These results are important for two reasons. First, they highlight that research intensity is influential for growth in India and, therefore, that the growth effects of R&D remain permanently positive. Second, the estimates show spillover effects are positive functions of the propensity to

[8] The estimated coefficient of the growth in R&D does not turn positive when the estimation period is limited to the pre-reform period of 1950-1990 (the results are not reported).

import and that the strongest benefits are derived from imports from R&D intensive countries. As the Indian economy has opened up along with the liberalization of the trade policies and the general globalization of the world economy over the past few decades, this result suggests that the government can explicitly increase growth by opening up the economy. Furthermore, India has increased its share of imports from R&D intensive countries since the 1980s.

Table 3: Parameter estimates of Eq. (5).

	$X = N$, $(X/Q)^d = N/L$, $(X/Q)^f = (R/Y)^f$		$X = N$, $(X/Q)^d = N/AL$, $(X/Q)^f = (R/Y)^f$		$X = R$, $X/Q = R/Y$		$X = R$, $X/Q = R/AL$		$X=PA$, $X/Q = PA/L$		$X = PG$, $X/Q = PG/L$	
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)
$\ln(A^{US}/A^{IN})$	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)
R^2	0.91		0.91		0.88		0.91		0.87		0.87	
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)
$\ln(A^{US}/A^{IN})$	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)
R^2	0.97		0.97		0.96		0.97		0.95		0.94	
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: 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 measures of growth in foreign research activity and foreign research intensity measures are weighted using import shares following the approach outlined in Eq. (6). LM-1 and LM-2 are Breusch-Godfrey LM statistics for no first- and second-order serial correlation in the residuals, respectively. Normality is the Jarque-Bera test of normal distributed residuals. Robust standard errors are obtained using the Newey-West procedure. A first-order autoregressive term is included in all regressions. The figures in round parenthesis are *t*-statistics and the figures in square brackets are *p*-values.

The estimated coefficients of the distance to the frontier ($\ln(A^{US}/A^{IN})$) are statistically and economically significant in all cases, thus providing convincing support to the conditional convergence hypothesis.[9] This finding suggests India's relative backwardness has played a

[9] We also investigated the interaction between domestic research intensity and distance to the frontier as predicted by the Schumpeterian growth model of Howitt (2000). However, the results do not give strong evidence in favor of this

favorable role in lifting its TFP growth during the sample period. It is interesting to note that the coefficient of distance to the frontier lies between 0.72 and 0.92, which is in the upper bound of the range of 0.30-0.80 found by Madsen (2008) for the OECD economies and slightly above the estimates of 0.65 by Lucas (2007) for the economies in the world classified as open. Given India's long distance to the technology frontier relative to non-US OECD economies, the estimates suggest that the growth-enhancing effects from technological gap are much stronger for India than non-US OECD economies. Finally, the level of financial liberalization has a statistically and an economically significant effect on TFP growth in most cases. The growth rate of financial liberalization, however, is found to have little impact on TFP growth.

A question is whether the model can account for the increase in the post-reform TFP growth rate. From the period 1960-1990 to 1991-2005, TFP growth rates increased by 1.1 percentage points. Simulating the model with the average coefficient estimates from the first two columns of the 10-year difference estimates in Table 3 yields the following results. Research intensity has contributed 2.8 percentage points to growth and research intensity spillovers have contributed 1.4 percentage points. Conversely, since the positive growth effects of the distance to the frontier have declined as technology gap has been reduced, the distance to the frontier has reduced growth rate by 1.7 percentage points. Similarly, financial liberalization has on average contributed to a 1.5 percentage points decline in growth because the financial liberalizations started from a very low level in 1991. The net effect of these factors is that the model predicts a one percentage point increase in the TFP growth rate which is very close to the actual increase in the TFP growth rate.

7. Robustness Checks

The surge in India's growth in the post-reform period raises the question of the extent to which the reforms have contributed to the Indian growth miracle and whether the results in the previous section are robust to the inclusion of control variables, particularly the variables that capture the most important economic reforms. Eq. (5) is extended with several control variables in this section.

The following control variables are considered. First, the ratio of foreign direct investment (FDI) to nominal GDP, which has increased substantially since the 1970s, particularly in the post-reform period. This variable is potentially important for growth because of the positive externalities associated with technologies transferred from countries that are often more technologically advanced than India. The estimation period starts in 1980 in the 10-year difference regressions containing FDI as a regressor because the data are first available from 1970. Second, patent

hypothesis. This may be an outcome of the high correlation between research intensity, distance to the frontier and their interaction. The results are reported and discussed in Appendix II.

protection is included in the regressions to allow for the possibility that intellectual property rights are conducive to an innovative environment. The patent protection index is an average of five different indicators that are constructed by Ginarte and Park (1997). Appendix V gives detail on the construction of this index. Third, the macro tariff rate, which is computed as the ratio of import duties and imports, is included as a proxy for trade barriers. Tariff rates increased steeply during the 1980s and have since been reduced to a level that is well below the level that has prevailed in the second half of the 20th century. Fourth, the interaction between the share of firms owned by non-residents (foreign firms) and the distance to the frontier is included in the regressions to examine the possibility that foreign firms facilitate the absorption of frontier technology.

Fifth, the share of total employment in agriculture is included in the regressions to cater for the growth effects of the transformation from agricultural dominance to manufacturing with higher productivity. Lucas (2007) argues that the share of employment in agriculture serves as a proxy for the educational attainment because workers in agriculture tend to be low skilled. Sixth, the interaction between the employment share in agriculture and the distance to the frontier is included as a control variable following the model of Lucas (2007). It is argued that the ability of countries to absorb technology that is developed at the frontier depends on the employment share in agriculture. Given the non-rival nature of technology developed at the frontier, the countries that possess adequate institutions and sufficiently educated labor force will be able to better absorb the frontier's technology. Since India meets the requirements for adequate institutions as defined by openness (Lucas, 2007), the interaction between employment share in agriculture and the distance to the frontier should influence growth negatively. Seventh, human capital is included as a control variable because along with R&D it is an essential knowledge variable in endogenous growth models.

Finally, the investment ratio is included in the estimates to allow for transitional dynamics (Jones, 1995a; Peretto, 1999; Howitt, 2000; Peretto and Smulders, 2002). The predictions of the endogenous growth models in this paper have been tested under the assumption that India has been growing along its balanced growth path. While this assumption may not be too strong in the cointegration estimates in which long-run relationships are estimated, it cannot be ruled out that the first-difference estimates have been affected by transitional dynamics.

The theoretical implications of transitional dynamics have been discussed in the semi-endogenous growth models of Jones (1995a), and in the Schumpeterian growth models of Peretto (1999) and Peretto and Smulders (2002). Transitional dynamics is important in two respects. First the parameter restrictions that distinguish semi-endogenous and fully-endogenous growth theories apply only in the steady-state equilibrium. Since India has experienced several shocks during the estimation period, the estimates may not have adequately captured the steady-state properties of the models. Second, investment does not play the same role of transitional dynamics in endogenous growth models as neoclassical models because neoclassical models focus on per capita output for any given exogenous TFP growth rate while endogenous growth models predominantly focus on TFP growth. However, Howitt (2000) shows in a Schumpeterian model that capital deepening induces more R&D by increasing the reward to innovation and by reducing the interest rate used for that reward. While these effects should be captured by the R&D terms, non-recorded R&D activities and other measurement errors, the transitional dynamics may not be adequately captured by the R&D terms. Furthermore, in the model of Peretto (1999) transitional dynamics is a function of capital intensity (capital per effective worker) in the intermediate sector.[10]

Table 4: Sensitivity checks for TFP growth estimates (10-year differences)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	-0.35	-2.06***	-2.59***	-2.18***	-2.81***	-2.24***	-2.25***	0.93	0.08	-9.35
$\Delta \ln X^d$	-0.25**	-0.24***	-0.25***	-0.23**	-0.22***	-0.17**	-0.20**	-0.21***	-0.03	-0.02
$\Delta mi \ln X^f$	0.53	0.20	0.09	-0.09	0.03	-0.15	-0.08	0.10	0.04	0.47
$\ln(X/Q)^d$	0.21**	0.21***	0.21***	0.17***	0.23***	0.17***	0.19***	0.30***	0.18***	0.87***
$mi \ln(X/Q)^f$	2.55*	5.00***	4.76***	5.02***	4.65***	4.69***	4.84***	6.18***	0.98	0.42
$\ln(A^{US}/A^{IN})$	0.99***	0.81***	0.86***	0.85***	0.86***	0.87***	0.87***	0.73***	0.85***	1.07***
$\ln FL$	-0.10	0.36***	0.39***	0.29***	0.32***	0.32***	0.32***	0.08*	-0.19*	0.41
$\Delta \ln FL$	-0.17***	-0.11**	-0.09	-0.09**	-0.10*	-0.10*	-0.10*	-0.13**	-0.15***	-0.11
ln(FDI/GDP)	0.04								0.05***	0.06**
ln(Patent protection)		-0.23***							0.25**	1.33***
ln(Macro tariff rate)			0.07							0.20
foreign firm share*lnDTF				-0.02						-0.01
ln(Agr emp share)					0.20					2.65**
(Agr emp share)*lnDTF						0.01				0.00
Δln(Investment/GDP)							-0.01		0.14**	0.17**
ln(Human capital)								-1.07		-1.15**
Δln(Human capital)								-0.32**	-0.37**	-0.32***
R^2	0.99	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.99	0.99
LM-1	2.73*	2.81*	2.51	4.34**	3.83*	3.04*	3.59*	2.02	8.13***	18.44***
LM-2	2.75	3.65	3.75	5.58*	5.69*	4.38	5.66*	2.47	9.73***	23.16***
Normality	0.54	5.63*	5.62*	2.89	4.94*	2.97	3.23	3.52	0.83	10.08***

Notes: Research intensity is measured as the fraction of R&D personnel in the total labor force. Estimation period covers 1960-2005, except for the estimates in columns (1), (9) and (10) in which the estimation period starts in 1980 (FDI data are first available from 1970). *, ** and *** denote 10%, 5% and 1% significance, respectively.

[10] During the transitional path, the specialization effect (driven by an increasing number of firms) dominates and TFP growth in the intermediate sector is a positive function of k (capital per effective worker) after k has grown sufficiently large. Eventually, the fragmentation effect, in which the number of small intermediate firms increases, dominates as k becomes bigger and the TFP growth rate in the intermediate sector converges to zero. After that period economy-wide TFP grows along its balanced growth path.

The regression results are displayed in Table 4. The model is regressed in ten-year differences and R&D is measured by the number of workers employed in the R&D sector. The control variables are added sequentially to Eq. (5) in the regressions of the first eight columns, while all the control variables are included in the estimates of the last column. Only the control variables that are significant and are of the right signs are included in the regressions in column 8.[11] The estimated coefficients of the growth in the investment ratio are significant in two out of three cases and are of the right sign, which suggests that the economy may not have been growing along its balanced growth path since 1950 (columns 7, 9 and 10). The estimated coefficients of the FDI-GDP ratio are significant in two of the three cases, which highlight the positive growth effects from opening up to foreign ownership (columns 1, 9, 10). Patent protection is significantly negative in one out of three cases, but significantly positive in the other two cases (columns 2, 9 and 10). The positive effect of strengthening patent protection framework found in the estimation period 1980-2005 is consistent with the significant increase in TFP growth rates observed in this period. It is, however, not clear how patent protection influences growth through its intended channel. Most theories suggest that patent protection increase innovative activity. Since innovative activity is already included in the regressions, the patent rights index may proxy other variables that are conducive to growth or account for some other non-measured innovations.

The estimated coefficients of levels and growth in human capital are either insignificant or significantly negative (columns 8, 9 and 10). These results are not surprising given that it has been difficult to find a robust relationship between growth and educational attainment in the literature (, 2001, 2006). The finding of no level effects from educational attainment is not surprising as Pritchett (2006) has pointed out that growth rates have not increased significantly over the past century in the OECD countries while secondary school enrolment rates have increased 25 fold during the same period. It is more surprising that growth has not been positively affected by educational attainment, which may reflect that educational attainment is measured with a large error. Regarding employment share in agriculture, it is puzzling that the estimated coefficients are positive and in one instance and even statistically significant (columns 5, 6 and 10). These variables are not included in the estimates in column (9) because they are likely to produce spurious results or simply reflect measurement errors.

Turning to the R&D related variables the estimated coefficients of $\Delta \ln X^{d}$ are consistently negative and are, in most cases, significant whereas the estimated coefficients of either $\ln(X/Q)^{d}$ or $mi\ln(X/Q)^{f}$ are, in most cases, highly significant and have the right positive signs. Thus, consistent with the finding in the previous sections, the results give support for Schumpeterian growth theory

[11] First differences of the control variables were also considered in the preliminary estimations. However, their estimated coefficients were all insignificant at the 10 percent level.

and no support for semi-endogenous growth theory. These results are important because they show that domestic and foreign research intensity remain important determinants of TFP growth even when variables representing the policy reforms are included in the estimates.

The results indicate that R&D intensity, R&D intensity spillovers, FDI/GDP and patent protection rights have permanent growth effects in most cases. The economy will continue growing as long as these variables remain positive. Importantly, all these variables are related to R&D and, as such, underscore that only knowledge variables can have permanent growth effects. Furthermore, the statistical properties of these variables are not inconsistent with Jones' (1995a) critique. None of these variables can increase beyond one (unlogged) and, therefore cannot increase continuously over time as levels R&D. In that sense, growth is bounded but permanent.

As a final check of the model and to investigate the effects of factor accumulation on growth, Eq. (5) is regressed using labor productivity as the dependent variable. The growth in the *K*-*Y* (capital-output) ratio and educational attainment are included as additional regressors, noting that the *K*-*Y* ratio is used instead of the capital stock to filter out the growth effects of TFP-induced productivity growth (see Madsen and Ang, 2009 for a discussion). The regressions give strong support for Schumpeterian models and no support for semi-endogenous growth (the results are reported in Appendix IV). Furthermore, the estimated coefficients of the *K*-*Y* ratio and educational attainment are either insignificant or negative and significant, regardless of whether total or private capital stock is used. The use of total or private investment ratios does not alter the results, suggesting that transitional dynamics may not have been an important factor in the Indian growth miracle.

8. Firm-level Analysis

The results in the previous section may have suffered from a small sample problem. More importantly the Schumpeterian results are derived under the assumption of free entry of firms. This assumption may not have been satisfied for India during the post-independence period given that the economy was subject to a set of industrial licensing requirements that restricted entry and expansion of both domestic and foreign firms, and this is dubbed “license raj” (Panagariya, 2002; Rodrik and Subramanian, 2005; Aghion et al., 2008). As discussed in Section 2, all licensing requirements were first fully withdrawn in 1991 and the cap on foreign direct investment in most industries was raised to 51 percent. Furthermore, as Laincz and Peretto (2006) argue, Schumpeterian growth models focus on the scale of a firm and not scale of 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.

8.1 Panel unit root and cointegration tests

Prior to examining whether the variables in Eqs. (3) and (4) are cointegrated, it would be useful to test for stationarity of the variables involved in the regressions. 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 adopt the Breitung's panel unit root procedure here. The results are reported in Table 5.

Table 5: Panel unit root tests

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: the null hypothesis is a unit root is present. 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 one percent level 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 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 stationary, 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 the predictions of Schumpeterian growth theory.

The Pedroni's (2004) panel cointegration procedure is used to test for panel cointegration. The optimal lag length is chosen to be one in all cases based on the AIC criterion. The results are reported in Table 6. Although $\ln R_{it}$ 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 non-negative and insignificant. However, there is robust support for Schumpeterian growth theory. All Pedroni's test statistics provide evidence in favor of the presence of cointegration between $\ln R_{it}$ and $\ln AL_{it}$ as well as $\ln R_{it}$ and $\ln Y_{it}$. More importantly, the second elements in the cointegrating vectors are statistically significant and have the right signs.

Table 6: Panel cointegration test

Model		Panel statistic		Group panel statistic		Cointegrating vector	
$\ln R_{it}$ and $\ln A_{it}$	*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)		
$\ln R_{it}$ and $\ln AL_{it}$	*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)		
$\ln R_{it}$ and $\ln Y_{it}$	*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: the null hypothesis of the test is that the variables are not cointegrated. An Intercept but no deterministic trend is included in the estimation. Figures in square brackets are *t*-statistics, and those in round parenthesis are *p*-values. *** indicates 1% levels of significance.

8.2 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:

$$\Delta \ln A_{it} = \varphi_0 + \varphi_1 \Delta \ln X_{it} + \varphi_2 \ln\left(\frac{X}{Q}\right)_{it} + \varphi_3 \ln\left(\frac{A^{\max_j}}{A}\right)_{it-1} + \varphi_4 \left(\frac{X}{Q}\right)_{it} \ln\left(\frac{A^{\max_j}}{A}\right)_{it-1} + TD + ID + \varepsilon_{it} \quad (7)$$

where $A^{max}{}_j$ is the industry technology frontier, *TD* is a vector of time dummies and *ID* is industry dummies,[12] 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 ($A^{max}{}_j$). The distance to the frontier and its interaction with research intensity is included following the predictions of the models of Howitt (2000) and Griffith *et al.* (2003, 2004). Unfortunately, import data by sources are not available at the firm-level, which preclude estimation of the effects of international technology spillovers through the channel of imports.

[12] These industries include Chemicals, Base Metals, Electrical Machinery, Non-Electrical Machinery, Electronics & Business Software, Food, Beverages & Tobacco, Textiles & Clothing, Plastic & Rubber, Paper, Transport, Wood Manufacturing, Furniture, Leather and Others.

Eq. (7) is estimated in 3- and 4-year differences using the fixed effects estimator.[13] Considering Table 7, it is evident that none of the estimates show a significant relationship between TFP growth and the growth in R&D ($\Delta \ln X_{it}$); thus providing no support for semi-endogenous growth theory. The regressions give support for Schumpeterian growth theory when research intensity is measured by the share of R&D expenditure in total output, but not when research intensity is measured as $(R/AL)_{it}$. These results are consistent with those of Laincz and Peretto (2006), who also find support for Schumpeterian growth theory but not semi-endogenous growth theory for the US. With regard to distance to the industry frontier ($\ln(A^{\max j}/A)_{i,t-1}$) and its interaction with research intensity, the results suggest that in all cases the former promotes TFP growth of a firm but not the latter. Thus, firms that are behind the technology frontier are able to use the technology that is developed elsewhere effectively and to close the gap between its TFP and the technological leader's TFP. However, given the statistical insignificance of the interaction term, we do not find evidence to support the hypothesis that firms behind the technology frontier will grow faster if they invest more in R&D.

Table 7: Panel TFP growth estimates

	3-year differences		4-year differences	
	$X/Q=R/Y$	$X/Q=R/AL$	$X/Q=R/Y$	$X/Q=R/AL$
Intercept	-0.418***	-0.623***	-0.005	-0.991***
	(-3.460)	(-5.950)	(-1.290)	(-5.470)
$\Delta \ln X_{it}$	0.002	0.002	-0.005	-0.005
	(1.040)	(0.950)	(-1.290)	(-1.370)
$\ln(X/Q)_{it}$	0.030*	0.008	0.080***	0.046
	(1.730)	(0.500)	(2.360)	(1.560)
$\ln(A^{\max j}/A)_{it-1}$	0.347***	0.338***	0.452***	0.439***
	(9.930)	(9.670)	(8.220)	(8.030)
$(X/Q)_{it}$ x $\ln(A^{\max j}/A)_{it-1}$	-0.154	0.001	0.082	0.001
	(-0.490)	(0.090)	(0.150)	(-0.380)
No. of obs.	1420	1420	985	985

Notes: the results are based on fixed effects estimations. The figures in round parenthesis are *t*-statistics. The number of observations is less than the cross-section and time-dimension of the data given that not all firms invest in R&D every year. * and *** indicate 10% and 1% levels of significance, respectively.

9. Conclusions and Implications of the Findings

The objective of this paper is two-fold: first, to test which of the two second-generation endogenous growth models is consistent with the data for India; and second, to examine the extent to which the roles R&D activity, international R&D spillovers, distance to the technology frontier

[13] Although GMM is more efficient than the OLS estimator in dynamic panels, the short time horizon in the *Prowess* data prevents us from using the GMM estimator.

and variables representing the economic reforms since independence, in explaining 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 1950-2005 and data for 590 manufacturing firms over the period 1993-2005, the results of the paper show 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. However, the estimates provide quite strong support for Schumpeterian growth theory. In particular, there exists an economically and statistically significant 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 increasing to infinity as predicted by the first-generation endogenous growth models. History and econometric tests suggest that TFP growth is stationary in the long run. 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 significant international R&D spillovers 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. These results point to the importance of foreign technology as one of the primary sources of India's TFP growth and the increasing TFP growth experienced in India during its post-reform period.

Although the findings point toward compelling evidence for Schumpeterian growth the following caveat is in order. The parameter restrictions that distinguish semi-endogenous and fully-endogenous growth theories apply only in the steady-state equilibrium. The Indian economy has not been in its steady state during the period 1950-2005. In fact, India has experienced a steady increase in the investment ratio, increasing R&D in proportion to GDP, and several policy reforms. The investment ratio and other variables were included in the estimates to allow for transitional dynamics. However, it is unlikely that these variables have captured all transitional dynamics and prevented them from influencing the key parameter estimates.

The paper, furthermore, finds that the post-reform growth spurt has been affected by the economic reforms that started in the 1980s and gained momentum after 1990. In fact, our results show that financial liberalization, increasing foreign direct investment in proportion of GDP, and patent protection rights have been influential for growth. Moreover, the economy has opened up as part of the reform programs and allowed a larger knowledge flow from abroad. Finally, the reforms have also given incitements to undertake R&D. For instance, weighted tax deductions and higher depreciation allowances associated with purchases of R&D-related machinery and materials have been granted by the government (Government of India, 2008). These results indicate that economic reforms are important in supporting a research driven economy.

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 TFP growth rates experienced in OECD economies in the post WWII period provided that it keeps its 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 at the frontier. This advantage will disappear as India approaches the technology frontier; however, this may take a long time.

Appendix I: 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 A1 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 A2, 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 A1: 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, the unit root tests suggest that the Indian data are more in line with the predictions of Schumpeterian growth theory than with semi-endogenous growth theory.

Table A2: 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.

Appendix II: The Growth Effects of Absorptive Capacity

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. The results in Table A3 suggest that, with only a few exceptions, that absorptive capacity is economically and statistically significant. This result is in line with the results of Griffith et al. (2004), Kneller and Stevens (2006) and Madsen (2008) who find 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 A3: 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.

Appendix III: Estimates with Trademarks as the Measure of Product Variety

Indirect measures of product variety have been used in the main text under the assumption that the number of products is equal to the size of the population in steady state. This appendix uses the number of trademarks applied for (T_t) as a direct measure of product varieties following Madsen (2008) (see also Mendonca et al., 2004; Mangani, 2007). More precisely, research intensity $(X/Q)_t$ is measured as $(N/T)_t$ and $(R/T)_t$.

First, consider Tables A4 and A5, which report the unit root and cointegration test results using trademarks as the measure of product variety. The unit root tests in Table A4 show that research intensity is stationary regardless of whether research intensity is measured in terms of R&D workers or real R&D expenditures, and whether the pre-reform period or the whole period is considered. These results are consistent with the predictions of Schumpeterian growth theory.

Table A4: Unit root tests for Schumpeterian growth models using trademarks as the measure of product variety

	Support for Schumpeterian?	
	Pre-reform (1950-1990)	Full sample (1950-2005)
$\ln(N/T)_t$	Yes	Yes
$\ln(R/T)_t$	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.

The results of regressing Eq. (4) using trademarks as a measure of product varieties are presented in Table A4. The estimates show that there is a significant cointegrated relationship between trademarks and the number of workers involved in R&D or the real R&D expenditure. The estimated coefficients trademarks in the cointegrating vectors are between -0.25 and -0.76, suggesting that the product variety elasticity of R&D is between 0.25 and 0.76. These elasticities are slightly smaller than the predictions the Schumpeterian theory, which suggests that the number of trademarks is increasing at higher rates than the number of product varieties.

The results of regressing Eq. (5) using trademarks as measures of product varieties are shown in Table A5. The results again give support for Schumpeterian growth theories and no support for semi-endogenous growth theory. The estimated coefficients of research intensity are positive and significant in five out of the eight cases. The estimated coefficients of growth in research activity are insignificant in six of the cases and negative and significant in the other two cases. Overall, these results are consistent with the estimates in the main text in which other measures of product varieties are used.

Table A5: Johansen cointegration tests for Schumpeterian growth theory (Eq. (4)) using trademarks as the measure of product variety

Model	Pre-reform (1950-1990) Trace statistic (Max-eigenvalue statistic)	Pre-reform (1950-1990) Cointegrating vector		Full sample (1950-2005) Trace statistic (Max-eigenvalue statistic)	Full sample (1950-2005) Cointegrating vector	
$\ln N_t$ and $\ln T_t$	19.219^{**} (17.747^{**})	1.000	-0.635^{***} [-23.136]	5.294 (3.661)	1.000	-0.385^{***} [-3.89]
$\ln R_t$ and $\ln T_t$	20.505^{***} (19.997^{***})	1.000	-0.764^{***} [-21.240]	19.069^{**} (18.681^{***})	1.000	-0.251^{***} [-3.061]

Notes: see notes to Table 2.

Table A6: Parameter estimates of Eq. (5) using trademarks as measures of product variety

	X = N $(X/Q)^d = N/TM$ $(X/Q)^f = (R/TM)^f$		*X = N* $(X/Q)^d = N/(A\text{x}TM)$ $(X/Q)^f = (R/TM)^f$		*X = R* $(X/Q)^d = R/TM$ $(X/Q)^f = (R/TM)^f$		*X = R* $(X/Q)^d = R/(A\text{x}TM)$ $(X/Q)^f = (R/TM)^f$	
	5-year differences							
Intercept	-3.91	(-4.60)	-4.85	(-3.54)	-3.96	(-2.14)	-3.06	(-1.44)
$\Delta \ln X^d$	0.06	(0.72)	-0.07	(-0.69)	-0.09	(-1.55)	-0.12	(-2.04)
$\Delta mi \ln X^f$	-0.15	(-0.46)	-0.19	(-0.61)	-0.21	(-0.56)	-0.18	(-0.53)
$\ln(X/Q)^d$	0.35	(2.77)	0.34	(2.22)	0.12	(1.18)	0.06	(0.60)
$mi \ln(X/Q)^f$	0.25	(0.88)	0.03	(0.11)	0.02	(0.21)	-0.02	(-0.24)
$\ln(A^{US}/A^{IN})$	0.78	(7.46)	0.75	(6.75)	0.84	(8.86)	0.83	(8.00)
$\ln FL$	0.43	(4.62)	0.44	(4.32)	0.35	(2.99)	0.28	(2.75)
$\Delta \ln FL$	0.04	(0.99)	0.04	(0.81)	0.06	(1.11)	0.06	(1.12)
ln(Trade openness)	0.50	(7.20)	0.54	(6.05)	0.36	(7.61)	0.34	(6.57)
R^2		0.91		0.91		0.91		0.90
LM-1	0.10	[0.75]	0.00	[0.96]	0.49	[0.49]	0.28	[0.59]
LM-2	0.11	[0.94]	0.15	[0.93]	0.71	[0.70]	0.66	[0.72]
Normality	3.44	[0.17]	6.99	[0.03]	0.81	[0.66]	0.99	[0.61]
	10-year differences							
Intercept	-5.13	(-9.37)	-7.52	(-4.48)	-6.15	(-2.38)	-3.84	(-1.12)
$\Delta \ln X^d$	0.10	(1.53)	-0.01	(-0.14)	-0.18	(-6.04)	-0.21	(-4.96)
$\Delta mi \ln X^f$	0.09	(0.30)	-0.14	(-0.53)	-0.03	(-0.11)	0.03	(0.10)
$\ln(X/Q)^d$	0.58	(8.75)	0.64	(4.08)	0.25	(2.16)	0.11	(0.83)
$mi \ln(X/Q)^f$	0.17	(0.79)	-0.03	(-0.08)	0.01	(0.08)	-0.06	(-0.50)
$\ln(A^{US}/A^{IN})$	0.81	(8.63)	0.83	(7.09)	0.92	(8.25)	0.90	(7.55)
$\ln FL$	0.56	(10.78)	0.66	(5.09)	0.54	(2.72)	0.38	(1.63)
$\Delta \ln FL$	-0.07	(-1.26)	-0.11	(-1.63)	-0.10	(-0.95)	-0.11	(-1.12)
ln(Trade openness)	0.77	(13.94)	0.98	(10.19)	0.62	(9.33)	0.61	(11.53)
R^2		0.97		0.96		0.97		0.97
LM-1	3.56	[0.06]	1.28	[0.26]	2.95	[0.09]	1.81	[0.18]
LM-2	5.93	[0.05]	1.56	[0.46]	4.74	[0.09]	2.89	[0.24]
Normality	2.68	[0.26]	2.36	[0.31]	16.96	[0.00]	16.15	[0.00]

Notes: *TM* refers to the number of trademarks applied. See notes to Table 3.

Appendix IV: Estimates using Labor Productivity Growth as the Dependent Variable

Growth has been explained in terms of TFP growth in the main body of the paper. Since per capita income growth equals TFP growth along the balanced growth path, TFP growth should be an adequate metric for growth. However, factor accumulation that is not TFP-induced may have been important for the Indian growth experience because the savings rate and educational attainment have been increasing during the estimation period. In other words, the Indian economy has been on a transitional path during the estimation period.

To incorporate non-TFP-induced factor accumulation into the analysis, consider the following homogenous Cobb-Douglas production function:

$$Y = K^{\alpha}(AhL)^{1-\alpha}, \tag{A1}$$

where L is raw labor. The production function can be written in the following convenient form:

$$\frac{Y}{L} = A\left(\frac{K}{Y}\right)^{\alpha/(1-\alpha)} h. \tag{A2}$$

Taking logs and differentiating yield the labor productivity growth rate, $g_{Y/L}$:

$$g_{Y/L} = g_A + \frac{\alpha}{1-\alpha} g_{K/Y} + g_h, \tag{A3}$$

which shows that labor productivity growth equals TFP growth along the balanced growth path, noting that (K/Y) and h are constant in steady state. In this formulation, labor productivity growth depends on the growth in the K-Y ratio and not the growth in K because the growth effects of TFP-induced capital deepening are filtered out. An increase in TFP increases the marginal productivity of capital and, consequently, results in capital deepening. Thus, the growth effects of the TFP-induced increase in capital stock are attributed to capital stock if the growth in capital stock is used in Eq. (A3) instead of the growth in the K-Y ratio, noting that the K-Y ratio is unaffected by TFP growth in steady state. Changes in the K-Y ratio along the balanced growth path are predominantly reflected by changes in savings rates.

The results of regressing Eq. (A3) combined with Eq. (5) are presented in Table A4. The estimated coefficients of growth in human capital are insignificant. However, it is hard to believe that increased educational standards have not increased labor productivity. It is more likely that the data are measured with errors. The estimated coefficients of the K^T-Y ratio are negative and only slightly significant in one of the two cases, where K^T is the sum of private and public capital stock.

This result may reflect that public capital stock has been the dominant part of the total capital stock and that public investment may have been unproductive. To investigate whether these results are driven by the productivity of public capital stock, the ratio of private capital stock (K^P) to output is used in the regressions in columns (2) and (6). The estimated coefficients of the K^P-Y ratio are negative and significant. As a final check on the possibility that capital deepening has affected Indian growth the investment ratio for private and total investment is included as a regressor. However, their estimated coefficients are insignificant in all four cases, suggesting that both capital deepening and transitional dynamics have not been influential for Indian growth (columns 3, 4, 7 and 8). These results reinforce the earlier findings that the Indian growth miracle has not been a result of factor accumulation and transitional dynamics.

Table A7: Estimates of per capita output growth

	$X = N; (X/Q)^d = N/L; (X/Q)^f = (R/Y)^f$				$X = R; X/Q = R/Y; (X/Q)^f = (R/Y)^f$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.49	0.45	0.56	0.56	1.05	1.07*	1.27*	1.10
$\Delta\ln X^d$	0.04	-0.06	0.06	0.05	0.03	0.03	0.00	0.02
$\Delta mi\ln X^f$	0.37	0.06	0.39	0.36	0.22	-0.04	0.35	0.26
$\ln(X/Q)^d$	0.05**	0.02	0.06**	0.05	0.21***	0.14***	0.18***	0.17**
$mi\ln(X/Q)^f$	1.20	5.38***	0.70	0.76	3.37**	5.12***	1.84	2.06*
$\ln(A^{US}/A^{IN})$	-0.07***	-0.04***	-0.08***	-0.08***	-0.06***	-0.06***	-0.08***	-0.08***
$\Delta\ln h$	0.01	-0.08	0.04	0.03	-0.03	-0.06	0.05	0.05
$\Delta\ln(K^T/Y)$	-0.03				-0.08*			
$\Delta\ln(K^P/Y)$		-0.36***				-0.30***		
$\Delta\ln(I^T/Y)$			0.01				-0.01	
$\Delta\ln(I^P/Y)$				-0.01				-0.01
$(X/Q)^d \times \ln(A^{US}/A^{IN})$	-0.37	-5.46	-1.48	1.14	-27.45	-23.34	-18.07	-18.96
$h \times \ln(A^{US}/A^{IN})$	0.15	-0.03	0.16	0.14	0.26	0.08	0.17	0.21
R^2	0.96	0.97	0.96	0.96	0.96	0.98	0.96	0.96
LM-1	3.25	3.17*	0.89	0.85	0.55	1.65	0.20	0.11
LM-2	3.18	4.76*	3.02	3.05	3.09	3.45	3.47	3.25
Normality	0.69	0.75	0.69	0.82	0.83	0.77	0.72	1.09

Notes: estimates are in 10-year differences. K^T is real total capita stock, K^P is real private capital stock, I^T is real total investment, and I^P is real private investment. *, ** and *** denote 10%, 5% and 1% levels of significance, respectively.

The TFP-related coefficient estimates are consistent with the TFP estimates in the main body of the text. Most of the estimated coefficients of research intensity are positive and significant while none of the coefficients of the growth in R&D are significant. Furthermore, the research intensity spillovers through the channel of imports are positive and significant in half of the cases while R&D growth spillovers are not in any of the cases. Finally, the estimated coefficients of the interaction between research intensity and the distance to the frontier and the interaction between educational attainment and the distance to the frontier are all statistically insignificant.

In summary, the estimates in this appendix have shown that: 1) factor accumulation has probably not played an important role in the Indian growth miracle; 2) labor productivity growth has predominantly been driven by R&D intensity; and 3) there is still strong support for Schumpeterian growth theories when labor productivity is used as the dependent variable.

Appendix V: Data and Measurement Issues

1) Aggregate data

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

$$TFP = \frac{Y_t}{K_t^{\alpha}(HL)_t^{1-\alpha}} \tag{A4}$$

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:

$$K_{t+1} = K_t(1-\delta) + I_{t+1} \tag{A5}$$

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:

$$K_0 = \frac{I_0}{\delta + g} \tag{A6}$$

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:

$$H = (1+r)^s \tag{A7}$$

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, patent applications by domestic residents, and number of trademarks applied. 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 and trademarks 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: