RELATING THE KNOWLEDGE PRODUCTION FUNCTION TO TOTAL FACTOR PRODUCTIVITY: AN ENDOGENOUS GROWTH PUZZLE

JEL Codes: O4, O3, and C5
Endogenous growth, knowledge spillovers, cointegration

June 2004

Yasser Abdih
Department of Economics
Bowling Green State University
Bowling Green, Ohio 43403
e-mail abdihy@cba.bgsu.edu

Frederick L. Joutz
Department of Economics
The George Washington University
Washington, DC 20052
e-mail bmark@gwu.edu

Abstract:

A central issue of endogenous growth theory has been the emergence of R&D based models of economic growth. These model(s) focus on the knowledge production function that describes the evolution of knowledge creation. The specification of the functions and the observable data has created debates over balanced growth paths and intertemporal knowledge spillovers. We link knowledge to economic growth by including total factor productivity in a system. Patent statistics are used to construct knowledge flows and stocks. The rate of production of new knowledge depends positively on the amount of labor engaged in R&D and the existing stock of knowledge available to these researchers. Time series evidence suggests that there are two long run cointegrating relationships. The first captures a long run knowledge production function; the second captures a long run positive relationship between TFP and the stock of knowledge (patents). The results indicate the presence of strong intertemporal knowledge spillovers. The long run elasticity of new knowledge creation with respect to the existing stock of knowledge is at least as large as unity. Second, the paper finds that the long run impact of the knowledge (patent) stock on TFP is small: doubling the stock of knowledge is estimated to increase TFP by only 10% in the long run. The paper provides evidence that the rate of diffusion of new knowledge into the productive sector of the U.S. economy has been slow over the past 20 years. This observation is consistent with the weak relationship found between the stock of knowledge and TFP. The paper interprets this evidence in light of the debate surrounding existing R&D-based growth models.

The authors would like to thank Vincy Fon, Tim Fuerst, Jim Hirabayashi, Arthur Goldsmith, Costas Mastrogianis, Stephanie Shipp, Holger Wolf, seminar participants at Washington and Lee University, and the Macro-International Workshop at George Washington University, for providing valuable comments. The first author would also like to thank John Kendrick. A preliminary version of this research was awarded The John Kendrick Prize, which supports research in the areas of productivity and growth.

1. INTRODUCTION AND CONTRIBUTION

The most recent advancement of endogenous growth theory has been the emergence of R&D-based models of growth in the seminal papers of Romer (1990), Grossman and Helpman (1991a, 1991b), and Aghion and Howitt (1992). This class of models aims to explain the role of technological progress in the growth process. R&D-based models view technology as the primary determinant of growth, and model it as an endogenous variable.

At the heart of R&D-based growth models is a knowledge/ technology production function that describes the evolution of knowledge creation. According to that function, the rate of production of new knowledge depends on the amount of labor engaged in R&D and the existing stock of knowledge available to these researchers. A crucial debate framed by Romer and Jones' work (within the R&D-based growth literature) is centered on the functional form of the knowledge production function. Specifically, the debate is centered on how strongly the flow of new knowledge depends on the existing stock of knowledge. Intuitively, the dependence of new knowledge on the existing stock is intended to capture an "intertemporal spillover of knowledge" to future researchers: knowledge or "ideas" discovered in the past may facilitate the discovery or creation of "ideas" in the present. Hence, the debate is concerned with the magnitude or the strength of these intertemporal knowledge spillovers. As will be discussed below, different assumptions on the magnitude of knowledge spillovers generate completely different predictions for long run growth.

This paper contributes to the empirical understanding of R&D-based growth models in the following ways. We use time series data for the U.S. economy over the postwar period and directly estimate the parameters of the knowledge production function. This allows us to directly assess the magnitude of knowledge spillovers, the source of the Romer-Jones debate. To achieve this goal, we exploit historical time series of patent filings to construct knowledge flows and stocks. Hence, the paper draws on an extensive body of work that uses patents as measures of innovative output and regards them as useful statistics for measuring economically valuable knowledge [e.g., Hausman, Hall, and Griliches (1984), Griliches (1989,1990), Joutz and Gardner (1996), and Kortum (1997)].We employ Johansen's (1988,1991) maximum likelihood cointegration procedure to estimate the U.S. knowledge production function. Cointegration techniques are needed because, like most macroeconomic time series, the inputs and output of the knowledge production function can be plausibly characterized as non-stationary and I(1) time series. Hence, if estimated using conventional methods like ordinary least squares, the knowledge production function will suffer from spurious correlations. The Johansen's cointegration procedure corrects for any spurious correlations that may exist amongst the inputs and output and explicitly accounts for the potential endogeneity of the inputs of the knowledge production function.

In his seminal paper, Romer (1990) assumes a knowledge production function where new knowledge is linear in the existing stock of knowledge, holding the amount of research labor constant. The implication of this strong form of knowledge spillovers is that the *growth rate* of the stock of knowledge is proportional to the amount of labor engaged in R&D. Hence, policies -such as subsidies to R&D- that increase the amount of labor allocated to research will increase the *growth rate* of the stock of knowledge. Since the Romer model is one in which long run per capita growth is driven by technological progress/ knowledge growth, such policies will increase long run per capita growth in the economy.

In an influential paper, Jones (1995a) questions the empirical validity of the Romer model. The Romer model predicts that an increase in the amount of research labor should increase the *growth rate* of the stock of knowledge, a prediction that depends critically on strong positive spillovers in knowledge production. Jones tests the validity of this prediction by appealing to data on total factor productivity growth (as a proxy for knowledge growth) and scientists and engineers engaged in R&D (as a proxy for research labor). He argues that in the U.S., the number of R&D scientists and engineers has increased sharply over the postwar period while total factor productivity (TFP) growth has been characterized by relative constancy at best. This weak relationship between the number of R&D scientists and engineers and TFP growth led Jones to conclude that the magnitude of knowledge spillovers assumed by Romer is too large. To be consistent with the empirical evidence, he argues that a smaller magnitude of knowledge spillovers needs to be imposed. Imposing a smaller magnitude of knowledge spillovers, however, alters the key implication of Romer's model. Specifically, in the modified model developed by Jones (1995b), long run growth depends only on exogenously given parameters, and hence is invariant to policy changes such as subsidies to R&D.

We study the cointegration properties of data on new knowledge (measured by the flow of new patents), the existing knowledge stock (measured by the patent stock), R&D scientists and engineers, and total factor productivity (TFP). We include TFP in the empirical model for three reasons. First, Jones (1995a) uses TFP as a measure of knowledge while we use the patent stock. Hence, the inclusion of TFP in the empirical model allows us to capture how closely our patent measure and Jones' measure are related. Second, long-run economic growth depends on total factor productivity which is the application and embodiment of knowledge. Third, it enables the estimated empirical model to shed some light on the observed weak relationship between TFP growth and the number of R&D scientists and engineers.

The paper finds two long run cointegrating relationships. The first captures a long run knowledge production function where the flow of new knowledge depends positively on the existing stock of knowledge and R&D scientists and engineers. The second captures a long run positive relationship between total factor productivity and the stock of knowledge (patents). The results indicate the presence of strong intertemporal knowledge spillovers, which is consistent with the Romer (1990) model. The long run elasticity of new knowledge creation with respect to the existing stock is at least as large as unity. However, the long run impact of the knowledge (patent) stock on TFP is small: doubling the stock of knowledge (patents) is estimated to increase TFP by only 10% in the long run. In other words, the results suggest that while R&D scientists and engineers greatly benefit from the knowledge and ideas discovered by prior research, the knowledge they produce seems to have had only a modest impact on measured total factor productivity.

These results seem to suggest a new interpretation of the empirical evidence documented by Jones (1995a). The observed weak relationship between the number of R&D scientists and engineers and TFP growth found by Jones is not necessarily an indication of weak intertemporal knowledge spillovers. We feel that knowledge, the output from researchers' effort, is an important intermediate step to TFP. The paper provides some evidence that the rate of diffusion of new knowledge into the productive sector of the U.S. economy has been slow over the past 20 years. The application and embodiment of knowledge into productivity is complex and diffuses slowly. Our empirical work contributes to understanding and reconciling some of the spillover effects and issues raised by Jones (1995a).

The rest of the paper is organized as follows. Section 2 presents a simple R&D-based growth model, with the focus on the Romer-Jones debate and the knowledge production function. Section 3 describes the data on the inputs and output of that function. Section 4 looks at the univariate and multivariate time series properties of the data and estimates the knowledge production function. Finally, section 5 offers some concluding remarks.

2. THE ROMER-JONES DEBATE ON KNOWLEDGE PRODUCTION

In this section, we present a simplified version of the R&D-based growth models of Romer (1990) and Jones (1995b). We focus on the basic elements and the key macroeconomic implications for long run growth. As such, we present the model in "reduced form" and hence suppress the micro-foundation and market structure components. This is done purely for ease of exposition.

2.1 A SIMPLE R&D- BASED GROWTH MODEL

The model has four variables: Output (Y), capital (K), labor (L), and technology or knowledge (A).[1] There are two sectors, a goods- sector that produces output, and an R&D sector that produces new knowledge. Labor can be freely allocated to either of the two

[1] In this paper, knowledge, technology, and "ideas" are used interchangeably.

sectors, to produce output (L_Y) or to produce new knowledge (L_A). Hence, the economy is subject to the following resource constraint $L_Y + L_A = L$.

Specifically, output is produced according to the following Cobb-Douglas production function with labor augmenting (Harrod-neutral) technological progress:

$$(1) \quad Y = K^{\alpha} (AL_Y)^{1-\alpha} \quad \text{, where } 0 < \alpha < 1$$

New knowledge or new "ideas" are generated in the R&D sector. Let A denote the stock of knowledge/technology available in the economy. The knowledge stock can simply be thought of as the accumulation of all the ideas that have been invented or developed by people. Then, $\dot{A}$ represents the flow of new knowledge or the number of new ideas generated in the economy at a point in time. New ideas are produced by researchers, L_A, according to the following production function:

$$(2) \quad \dot{A} = \bar{\delta}\, L_A$$

where $\bar{\delta}$ denotes (average) research productivity, i.e. the number of new ideas generated per researcher. $\bar{\delta}$, in turn, is modeled as a function of the existing stock of knowledge/ideas (A) and the number of researchers (L_A) according to:

$$(3) \quad \bar{\delta} = \delta\, A^{\phi}\, L_A^{\lambda-1} \ , \quad \delta > 0$$

where δ, φ, and λ are constant parameters. The presence of the term A^{ϕ} in (3) is intended to capture the dependence of current research productivity on the stock of ideas that have already been discovered. Ideas in the past may facilitate the discovery or creation of ideas in the present, in which case current research productivity is increasing in the stock of knowledge ($\phi > 0$). To quote Jones (1995b), "The discovery of calculus, the invention of the laser, and the creation of semiconductors are all examples of ideas that most likely raised the productivity of the scientists who followed." Hence, $\phi > 0$ captures a positive “spillover of knowledge” to future researchers and is referred to as the “standing on shoulders” effect. Alternatively, it is possible that the most obvious ideas are discovered first and new ideas become increasingly harder to find over time. In this case, current research productivity is decreasing in the stock of ideas already discovered. This corresponds to $\phi<0$, the “fishing out” effect. [2]

The presence of the term $L_A^{\lambda-1}$ in (3) captures the dependence of research productivity on the number of people seeking out new ideas at a point in time. For example, it is quite possible that the larger the number of people searching for ideas is, the more likely it is that duplication or overlap in research would occur. In that case, if we double the number of researchers (L_A), we may less than double the number of unique ideas or discoveries ($\dot{A}$). This notion of duplication in research or the “stepping on toes” effect can be captured mathematically by allowing for $0<\lambda<1$, in which case research productivity is decreasing in L_A.

Taken together, equations (2) and (3) suggest the following knowledge or “ideas” production function:

$$(4) \quad \dot{A} = \delta \; L_A^{\lambda} A^{\phi}$$

That is, the number of new ideas or new knowledge at any given point in time depends on the number of researchers and the existing stock of ideas.

2.2 GROWTH IMPLICATIONS OF THE MODEL

Given the above setup, it can be easily shown that there exists a balanced growth path / steady state for this economy, defined as a situation in which all variables grow at constant (possibly zero) rates. Along this path, output per worker (y) and the capital–labor ratio (k) grow at the same rate as that of technology (A):

[2] The case where $\phi=0$ allows the “fishing out” effect to completely offset the “standing on shoulders” effect. That is, current research productivity is independent of the stock of knowledge.

$$(5) \quad g_y = g_k = g_A$$

where g_y, g_k, and g_A respectively denote the steady state growth rate of y, k, and A. Hence, R&D based growth models share the prediction of the neoclassical Solow model that technological progress is the source of sustained per capita growth. If technological progress ceases, so will long run per capita growth. Therefore, to solve for the steady state per capita growth rate in this economy, it suffices to solve for g_A, which is in turn determined by the knowledge production function as shown below. We focus on two versions of that function: Romer (1990) and Jones (1995b). As will be discussed below, their versions have completely different implications for long run growth. Those implications depend critically on the magnitude of the knowledge spillover parameter assumed (ϕ in equation 4).

2.3 THE ROMER (1990) MODEL

Romer (1990) assumes a particular form of the knowledge production function in (4). He imposes the restrictions ϕ=1 and λ=1. The key restriction made by Romer, however, is ϕ=1. This makes $\dot{A}$ linear in A, and hence generates growth in the stock of knowledge ($\dot{A}$/A) that depends on L_A unit homogeneously:

$$(6) \quad \frac{\dot{A}}{A} = d\,L_A$$

Equation (6) pins down the steady state growth rate of the stock of knowledge, g_A, as

$$(7) \quad g_A = d\,L_A$$

That is, the steady state growth rate of the stock of knowledge (and per capita output by equation 5) depends positively on the amount of labor devoted to R&D. This key result has important policy implications: Policies which permanently increase the amount of labor devoted to R&D- a subsidy that encourages research for example - have a permanent long run effect on the growth rate of the economy. This “growth effects” result is a hallmark of the Romer (1990) model and many existing R&D-based endogenous growth models, including the important contributions of Grossman and Helpman (1991a, 1991b) and Aghion and Howit (1992). This result stands in sharp contrast to the neoclassical Solow model, in which changes in variables that are potentially affected by policy have short-run/medium-run effects but no long run growth effects.

2.4 THE JONES (1995 a) CRITIQUE

Equation (7) predicts "scale effects": an increase in the *level* of resources devoted to R&D- as measured by L_A – leads to an increase in the *growth rate* of the economy. This "scale effects" prediction of the Romer model is rooted in the knowledge production function (equation 6), which states that technological growth should be proportional to the number of research workers. In a very influential paper, Jones (1995a) presents time series evidence against scale effects using a measure for L_A and one for $\dot{A}/A$ for the U.S. over the postwar period. He represents L_A by the number of scientists and engineers engaged in R&D. This is perfectly reasonable, since theoretically, L_A captures the R&D workforce. Jones uses Total Factor Productivity (TFP) growth as a proxy for $\dot{A}/A$, which is shown in figure 1a for the US economy. The pattern of TFP growth is well known: TFP growth appears to fluctuate around a relatively constant mean of about 1.4% per year over the postwar period. Therefore, L_A should, like $\dot{A}/A$, be relatively constant and exhibit no persistent increase. Otherwise, the Romer's knowledge production function and the resulting scale effects are inconsistent with the time series evidence.

Figure 1b plots $L_{A,}$ as measured by the number of scientists and engineers engaged in R&D for the U.S. economy. As Figure 1b reveals, L_A is not relatively constant over the postwar period. Rather, it exhibits a very strong upward trend, rising from about 100,000 in 1950 to about 1 million by 1997. Therefore, the knowledge production function in equation (6), which lies at the heart of the Romer (1990) model, is inconsistent with the time series data. It is important to emphasize that the criticism by Jones is not exclusive to the Romer (1990) model, but rather is a criticism against many existing R&D-based endogenous growth models that share Romer's knowledge production function.

2.5 THE JONES' (1995b) ALTERNATIVE

Since the rejection of the scale effects prediction is rooted in the incongruence of the knowledge production function with the time series data, it seemed sensible for Jones to tackle and modify its functional form in an attempt to come up with an alternative specification that is consistent with the observed time series pattern of the data.

Jones (1995b) actually shows that relaxing the assumption ϕ=1 generates a steady state that is consistent with the rising number of research workers observed in the data.

To do that, consider once again the modified knowledge production function in (4) and divide both sides of that equation by A to get:

$$(8) \quad \frac{\dot{A}}{A} = d \; \frac{L_A^{l}}{A^{1-f}}$$

In the steady state, the growth rate of A is constant by definition. Therefore, the right-hand-side of (8) must be constant in the steady state, which means that L_A^{l} and A^{1-f} must grow at the same rate. That is,

$$(9) \quad l \, \frac{\dot{L}_A}{L_A} = (1-f)\frac{\dot{A}}{A}$$

Now, l is a positive parameter and $\dot{A}$/A is always positive and constant in the steady state. Therefore, (9) implies that a constant steady state growth of A will be consistent with a rising L_A, i.e. $\dot{L}_A$/ L_A>0, provided that ϕ is less than unity. Stated slightly differently, imposing $\phi < 1$ guarantees a steady state in the presence of a rising number of research workers. Hence, Jones (1995b) argues, assuming $\phi < 1$ is consistent with the observed relative constancy of TFP growth (the proxy of $\dot{A}$/A used by Jones) in spite of the rising trend of R&D scientists and engineers. Moreover, with $\phi < 1$ imposed, the scale effects of the Romer (1990) model are removed. This can be seen formally by solving for the steady state growth rate of A from equation (9) as:

$$(10) \quad g_A = \frac{l}{1-f} \frac{\dot{L}_A}{L_A}$$

That is, the long run *growth rate* of the stock of knowledge [which is also the long run growth rate of per capita output by (5)] depends on the *growth rate* of L_A rather than its *level*. Note that positive knowledge spillovers are not ruled out. The parameter capturing knowledge spillovers, ϕ, may plausibly be positive and large, Jones argues. What the above discussion does suggest is that the *degree* of positive knowledge spillovers assumed by Romer is arbitrary and is inconsistent with the time series evidence. A weaker magnitude of such spillovers is needed to achieve congruency with the evidence.

Now, along the balanced growth path/steady state, the growth in the number of research workers will be equal to the growth rate of the labor force/population. If it was greater, then the number of researchers will eventually exceed the labor force, which is not feasible. Let n denote the growth rate of the labor force/population, which Jones (1995), following the literature, assumes to be exogenously given. Then the above argument implies that in the steady state, $\dot{L}_A / L_A = \dot{L} / L = n$. Substituting this relationship into (10) yields

$$(11) \quad g_A = \frac{l}{1-f} n$$

Two important features of equation (11) are worth noting. First, one does not want to interpret it in a cross-country framework as saying that countries with a higher population growth enjoy faster growth. Rather, the model should be thought of as describing the evolution over time of a country at the forefront of the technological frontier, a country like the US. That said, what is the intuition underlying this equation? The intuition is best seen by assuming λ =1 and ϕ >0, in which case the knowledge production function reduces to $\dot{A} = d \;\; L_A \; A^f$, where positive knowledge spillovers are assumed. At any given point in time, researchers draw on the existing stock of knowledge to create new knowledge. New knowledge, in turn, adds to the stock and the latter feeds back into subsequent new knowledge through a positive knowledge spillover effect (ϕ >0). Over time, both new knowledge ($\dot{A}$) and the stock A are growing; but since the returns to knowledge accumulation is less than unity (ϕ <1), the ratio of new knowledge ($\dot{A}$) to the stock A will be falling if the number of researchers is constant. To offset that potential fall in $\dot{A}/A$, the number of researchers must increase over time- because of population growth for example- sustaining growth in the model. This explains the dependence of the long run value of $\dot{A}/A$ on n in equation (11).

Second, equation (11) implies that long run growth depends on ϕ, λ, and n, parameters that are usually assumed to be exogenously given! Hence, long run growth in the Jones'(1995b) model is independent of policy changes such as subsidies to R&D. Because the returns to knowledge accumulation is assumed to be less than unity (ϕ <1), such changes will affect the growth of A along the transition path to a new steady state, and these "transitional growth effects" will be translated into long run *level* effects. Simply stated, subsidies to R&D will alter the long run *level* of the stock of knowledge but not its long run growth rate. In the modified model of Jones (1995b), long run growth is invariant to policy. As Jones (1995b) concludes his paper "Nothing in the U.S. experience appears to have had a permanent effect on growth. In light of this evidence, the invariance result maybe exactly what the data requires."

3. THE DATA

In this section we describe the variables used in empirically reconsidering the theoretical relationships between the knowledge production function and productivity. The four include: patent applications, the stock of patents, the number of scientists and engineers engaged in R&D, and total factor productivity. The sample frequency is annual and is available from 1948 to 1997. Variables in levels will be transformed into natural logarithms.

Patent applications serve as a valuable resource for measuring innovative activity and have been extensively used in the patent literature as measures of technological change [see, e.g., Hausman, Hall, and Griliches (1984), and Kortum (1997)]. Also, Griliches (1989, 1990) and Joutz and Gardner (1996) argue that patent applications are a good measure of technological output. Firms have invested resources in developing a new technology, which they feel has economic value and they are willing to submit an application to capture rents from their initial investments. As such, this paper follows the patent literature and uses patent applications to construct knowledge flows and stocks.[3]

The output of the knowledge production function should reflect new knowledge created by US researchers. As such, we use domestic patent applications (DP) filed at the U.S. Patent and Trademark Office (USPTO) to measure new knowledge. The USPTO provides information on the number of patent applications filed from 1840 to present. These include patents for invention, designs, and plants. This data is available on-line from the US Patent and Trademark Office web site, at:

http://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm

Figure 2 plots the log of domestic patent applications (dp). There appears to be an overall upward trend in the series over the entire sample period 1948-1997. In fact, domestic patent applications grew at an average annual rate of 1.7% between 1948 and 1997. Looking at sub-periods, domestic patent applications seem to have grown fairly rapidly during the 1950s till about the mid 1960s. The average annual growth rate was about 1.6% between 1948 and 1965. However, from the mid 1960s till about the mid 1980s,

[3] Note that we measure knowledge/technology using patent applications rather than patent grants. The lag between application and grants could be quite long and it varies over time partly due to changes in the availability of resources to the U.S. Patent Office. This notion is best articulated by Griliches (1990): "A change in the resources of the patent office or in its efficiency will introduce changes in the lag structure of grants behind applications, and may produce a rather misleading picture of the underlying trends. In particular, the decline in the number of patents granted in the 1970s is almost entirely an artifact, induced by fluctuations in the Patent Office, culminating in the sharp dip in 1979 due to the absence of budget for printing the approved patents." This paper views patent applications as a much better measure of knowledge/ technology than patent grants. Also, It is widely believed that patent application data is a better measure of new knowledge produced in an economy than R&D expenditures [see, e.g., Joutz and Gardner (1996)]. The reason is that R&D expenditures are more properly thought of as inputs to technological change while patents are an output. Hence patent applications more closely approximate the output of the knowledge production function in R&D-based growth models than R&D expenditures.

domestic patent applications seem to be characterized by relative constancy. What is particularly striking in figure 2 though is the behavior of the series since the mid 1980s: since about 1985, domestic patent applications have increased dramatically at an average annual rate of 5.1%, an increase substantially higher than any other witnessed over the entire course of US history. There are several hypotheses in the literature that attempt to explain this phenomenon. It is important to put the magnitude of the increase and the informational content in perspective.

Kortum (1997) and Kortum and Lerner (1998) articulate that the sharp increase in patenting witnessed since the mid 1980s is the result of a major institutional change in the US patent policy, a change that substantially benefited patent holders. Specifically, in 1982, Congress established the Court of Appeals of the Federal Circuit (CAFC), a specialized court to hear patent cases. (Patent appeals cases were heard before district courts. Before 1982, the treatment of appeals varied by district.) The court unified and made alterations to the patent doctrine with the purpose of enhancing the efficiency of the patent system. It also seemed to have broadened the rights of patentees as its decisions and rulings have been widely conceived as being "pro-patent". This institutional change and the stronger level of patent protection it brought could explain why patenting has surged since 1985.

Alternatively, and more interestingly, the sharp increase in patenting might reflect a "real" surge in discovery and innovation. Patents have long been used and are used in this study as an outcome of innovative activity and technological change.

Several pieces of evidence in the literature are supportive of this view. First, Greenwood and Yorukoglu (1997) document that the 1980s and 1990s witnessed "an explosion of formation of new firms and innovation in the high-tech industries, particularly in the information technology, biotechnology, and software industries." Hence, the sharp increase in patenting may indicate a "technological revolution" as emphasized by these authors.

Second, it is quite possible that the use of information technology itself in the discovery of new ideas might have substantially boosted research productivity. Arora and Cambardella (1994) argue that this was an important source of accelerating technological change.

A third possibility, emphasized by Kortum and Lerner (1998), is that the sharp increase in patenting since the mid 1980s indicates an increase in innovation driven by improvements in the management of R&D. In particular, there has been a reallocation of resources from basic research toward more applied activities and hence a resulting surge in patentable discoveries. As Kortum and Lerner [p. 287] point out "Firms are restructuring, redirecting and resizing their research organizations as part of a corporate-wide emphasis on the timely and profitable commercialization of inventions combined with the rapid and continuing improvement of technologies in use."

The stock of knowledge is derived from the cumulated number of total patents applied for by both US and foreign inventors. Patent filings are converted into a stock measure (STP) using the perpetual inventory method with a depreciation rate of 15% [see Appendix A for details on data measurement].

This is typical in the U.S. patent literature [e.g., Griliches (1989), Joutz and Gardner (1996)]. While this approach is ad-hoc and not necessarily justified by theory, researchers have typically checked the robustness of their results to changes in the depreciation rate. We experimented with constructing stocks using 0,5, and 10 % depreciation rates, and found that the precise rate made very little difference. Hence, the results presented in this paper are not sensitive to changes in the depreciation rate on the stock. Also, as will be shown below, the model estimated in the paper is stable over the sample period. This is particularly important since it implies that, among other things, the long run impact of the existing stock of knowledge on the flow of new knowledge has remained stable over time.

Figure 3 plots the log of the stock of total patent applications (stp). There appears to be a strong upward trend in the series over the entire sample 1948-1997. In fact, between 1948 and 1997, the stock of total patent applications grew at an average annual rate of 1.9%. There also appears to be a substantially stronger trend since the mid 1980s: prior to 1985, the stock grew at an average annual rate of 1.2%. However, after 1985, the average annual growth rate of the stock more than tripled to about 4%. This more rapid increase in the stock since the mid 1980s captures the more rapid increase in (the number of) domestic patent applications that occurred over that period.[4]

Figure 4 plots the log of the total number of scientists and engineers engaged in R&D activities (s&e) in the United States. This measure was used by Jones (2002) and represents scientists and engineers employed in industry, the federal government, educational institutions, and nonprofit organizations. It is accepted as the best proxy for the primary input or effort in the knowledge production process. The data for the period 1979-1997 is obtained from the National Science Foundation, *Science and Engineering Indicators- 2000*. This source is available online at: http://www.nsf.gov/sbe/srs/seind00/start.htm. For years prior to 1979, the data is taken from Jones (2002) and Machlup (1962) who in turn obtain their data from the National Science Foundation [NSF (1993, 1962, 1961, and 1955)].

The series exhibits a very strong upward trend over the past 50 years. In the late 1960s through the early 1970s, however, employment of R&D scientists and engineers seems to have declined. The National Science Foundation [NSF (1998)] documents that this is probably due to the fact that both the federal government and the business sector de-emphasized funding for certain research programs during that period. In particular, federal funding for space related R&D declined substantially in the late 1960s and early

[4] By definition, the stock of total patent applications includes (1) the cumulated numbers of domestic patent applications (i.e., patent applications by US inventors at the US patent office), and (2) the cumulated numbers of foreign patent applications (i.e., patent applications by foreign inventors at the US patent office). We examined the data on the number of foreign patent applications, and found that it is characterized by a sustained smooth rise over the period 1948-1997. There is no evidence of a more rapid increase in the series beginning in the mid 1980s. As such, the more rapid increase in the stock of total patent applications since the mid 1980s indeed captures the more rapid increase in the numbers of domestic patent applications.

1970s after the thrust of funding in the early to mid 1960s, during which the US had invested substantial resources in the “space race” with the Soviet Union. Overall, however, the number of R&D scientists and engineers grew substantially at an average annual rate of 4.3% over the period 1948-1997.

Total factor productivity (TFP) for the private business sector of the U.S. economy was obtained from Larry Rosenblum at the Office of Productivity and Technology, the Bureau of Labor Statistics, U.S. Department of Labor. Figure 5 shows the plot for the log of total factor productivity (tfp); it follows an upward trend over the postwar era. In fact, it grew at an average annual rate of about 1.4% between 1948 and 1997. As shown in figure 5, total factor productivity growth appears to have slowed since 1973- the well-known productivity slowdown. Prior to 1973, the average annual growth rate of total factor productivity was 2.1%. After 1973, the average annual growth rate declined to about 0.7%.

4. ESTIMATION OF KNOWLEDGE PRODUCTION FUNCTIONS

We employ the general-to-specific modeling approach advocated by Hendry (1986). It attempts to characterize the properties of the sample data in simple parametric relationships which remain reasonably constant over time, account for the findings of previous models, and are interpretable in an economic sense. Rather than using econometrics to illustrate theory, the goal is to "discover" which alternative theoretical views are tenable and test them scientifically.

The approach begins with a general hypothesis about the relevant explanatory variables and dynamic process (i.e. the lag structure of the model). The general hypothesis should be considered acceptable to all adversaries. Then the model is narrowed down by testing for simplifications or restrictions on the general model.

The four macroeconomic and innovation variables are linked through two main relationships. The long-run knowledge production function and the long run relationship between total factor productivity and the stock of total patents (knowledge) can be specified as:

$$\begin{aligned} dp &= F(stp, s\&e) \\ tfp &= G(stp) \end{aligned}$$

where lower-case letters denote variables in natural logarithms. That is, dp denotes the log of the number of domestic patent applications, stp denotes the log of the stock of total patent applications, and s&e denotes the log of the number of scientists and engineers engaged in R&D. According to the above production function, US R&D scientists and engineers produce US patents, but they draw upon the "world" stock of knowledge.[5] Also, the function F(.) is assumed to be linear.[6]

Since Jones (1995a) used total factor productivity as a measure of knowledge, we also include the relation G(.) for the log of total factor productivity, tfp, in the model. This allows us to capture how closely our patent measure and Jones' measure are related, and allows us to interpret the results in terms of Jones' time series evidence on total factor productivity and R&D scientists and engineers. We look at the transmission mechanism by separating the R&D effort and output. The total stock of patents represents the cumulative R&D output which leads to higher productivity. This is consistent with the substantial micro productivity literature [Jaffe, Henderson, and Trajtenberg (1993) and Thompson and Kean (2004)] that postulates a positive dependence of total factor productivity on the stock of patents.

The first step in the modeling approach examines the time series properties of the individual data series. We look at patterns and trends in the data and test for stationarity and the order of integration. Second, we form a Vector Autoregressive Regression (VAR) system. This step involves testing for the appropriate lag length of the system, including residual diagnostic tests and tests for model/system stability. Third, we test the system

[5] Below, we also use the stock of domestic patents as an alternative measure of the stock of knowledge. We compare the results from using such a measure with the results where the stock of total (domestic and foreign) patents is used.

[6] Recall that the R&D-based growth models of Romer (1990) and Jones (1995b) assume a Cobb-Douglas specification for the knowledge production function expressed in terms of the levels of the variables. Since the function F(.) in the text is expressed in terms of the log-levels of the variables, it is therefore assumed to be linear.

for potential cointegration relationship(s). Data series integrated of the same order may be combined to form economically meaningful series that are integrated of lower order. Fourth, we interpret the cointegrating relations and test for weak exogeneity. Based on these results a conditional error correction model of the endogenous variables may be specified, further reduction tests are performed and economic hypotheses tested. This last step will not be preformed, because the primary goal is to understand the long-run relationships.

4.1 INTEGRATION ANALYSIS

Figures 2-5 showed significant trends in the series and the autocorrelations were quite strong and persistent. Nelson and Plosser (1982) found that many macroeconomic and aggregate level series are shown to be well modeled as stochastic trends, i.e. integrated of order one, or I(1). Simple first differencing of the data will remove the non-stationarity problem, but with a loss of generality regarding the long run "equilibrium" relationships among the variables. We performed the standard augmented ADF test in both levels and differences with a constant and trend. Table 1 contains the results in five columns and is divided in two. The top half is for the tests in levels and the bottom is for the tests in first differences or whether the series are I(1) and not I(2) respectively. The first column lists the variables. The Akaike information criterion was used to set the appropriate lag-length for the dependent variable in each test and is provided in the second and fourth columns. The t-ADF statistics are reported in the third and fifth column.

We cannot reject the null hypothesis of a unit root for all four variables in levels.[7]
Domestic patents, total factor productivity, and scientists and engineers engaged in R&D reject the null of a unit root in first differences while the stock of patents does not. However, a recursive analysis of the coefficient estimate and the t-ADF suggest that it is non-constant with a break right where one might expect, 1985. In our preliminary look at

[7] Table 1 also includes a fifth variable, sdp. sdp is the stock of domestic patents constructed from the number of domestic patent applications using the perpetual inventory method. This variable will be discussed at the appropriate time below. For now, it suffices to say that the statistical arguments mentioned in the text for stp also apply to sdp.

the data we saw the acceleration in the propensity to patent and its impact on the stock of patents. Figure 6 presents a plot of the first difference in the logarithm of the patent stock measure. It appears there is a permanent shift in the mean starting the in the mid 1980s. The Perron (1989) structural break procedure was used to test for whether there was mean shift in the first difference process which caused the I(1) findings. We could not reject the hypothesis of an I(1) process after correcting for the (structural) mean shift.

4.2 COINTEGRATION ANALYSIS

Our analysis of the inputs and outputs of knowledge production suggest that the processes are non-stationary.[8] This has implications with respect to the appropriate statistical methodology. While focusing on changes in knowledge production eliminates the problem of spurious regressions, it also results in a potential loss of information on the long-run interaction of variables (e.g., Davidson, Hendry, Srba, and Yao (1978)). We examine the hypothesis of whether there exist economically meaningful linear combinations of the I(1) series: (domestic) patent filings, the stock of patents, scientists and engineers engaged in R&D, and total factor productivity that are stationary or I(0). The Johansen maximum likelihood procedure is used for the analysis. The procedure begins with specifying a VAR system

(12) $$Y_t = p_0 + \sum_{i=1}^{p} p_i Y_{t-i} + \Psi D_t + e_t$$

where

$$Y_t = \begin{bmatrix} \textit{Patent Filings}_t \\ \textit{Patent Stock}_t \\ \textit{R \& D Scientists \& Engineer}_t \\ \textit{Total Factor Productivity}_t \end{bmatrix}, \quad e_t \sim IN(0, \Omega), \; and \quad D_t = \begin{bmatrix} Stepdum86_t \\ Impulse9495_t \\ Impulse96_t \\ Trend_t \end{bmatrix}.$$

---

[8] Elliott (1998) points out the pitfalls in using cointegration methods when the data are stationary. However, this is a limitation of any research using this methodology.

Y_t is $(n \times 1)$ and the p_i's are $(n \times n)$ matrices of coefficients on lags of Y_t. D_t is a vector of deterministic variables that can contain a linear trend, dummy-type variables, or other regressors considered to be fixed and non-stochastic. Finally, e_t is a $(n \times 1)$ vector of independent and identically distributed errors assumed to be normal with zero mean and covariance matrix W [i.e., e_t ~ i.i.d. N(**0**, W)]. As such, the VAR comprises a system of n=4 equations, where the right-hand side of each equation comprises a common set of lagged and deterministic regressors.

The VAR includes our four series: the log of domestic patent applications, dp, the log of total factor productivity, tfp, the log of the (lagged) stock of total patent applications, stpl1[9], and the log of the number of scientists and engineers engaged in R&D, s&e. The VAR also includes a constant, a trend term, and three dummy variables. The first dummy variable is Stepdum86, which takes the value of one after 1985 and zero otherwise. The inclusion of this variable is intended to capture the dramatic increase in patenting since the mid 1980s as discussed in detail above. The second dummy variable is Impulse9495 which takes the value of one in 1994 and 1995 and zero otherwise, and the third is Impulse96, which is zero except for unity in 1996. Economically, Impulse9495 captures several institutional changes in the US patent policy: the movement towards the typical international patent system policy of granting 20 awards instead of 17 year awards, twelve-year patent renewal fees were collected for the first time in the United States in 1994/1995 [Kortum 1997], whereas Impulse96 captures the instantaneous negative response by agents facing an increased cost of patent applications [see Figure 2]. Statistically, the residuals of the VAR fitted without Impulse9495 and Impulse96 have large outliers in 1994, 1995, and 1996 and induce misspecification in the residuals. This problem is resolved by the inclusion of the two dummy variables, which results in a substantial improvement in the fit of the model and much better residual diagnostics. In addition, the inclusion of the three dummies Stepdum86, Impulse9495 and Impulse96 ensures a statistically stable/constant VAR as will be shown below.

[9] stpl1 is simply stp lagged one period. Since stp is calculated as end of period stocks, we enter it with a lag in the VAR and cointegration analysis.

Following Johansen and Juselius (1990), the VAR model provides the basis for cointegration analysis. Adding and subtracting various lags of Y yields an expression for the VAR in first differences. That is

$$(13) \qquad \Delta Y_t = p_0 + p \; Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \Psi \; D_t + e_t$$

where $\Gamma_i \; = -(p_{i+1} + ... + p_p), \quad i = 1,...,p-1 \quad and \quad p \equiv \left(\sum_{i=1}^{p} p_i \right) - I$

The VAR model in differences is actually a multivariate form of the ADF unit root test. If p is a zero matrix, then modeling in first differences is appropriate. The matrix p may be of full rank or less than full rank, but of rank greater than zero. When $rank(p) = n$, then the original series are not I(1), but in fact I(0); modeling in differences is unnecessary. But, if $0 < rank(p) \equiv r < n$, then the matrix p can be expressed as the outer product of two full column rank (nxr) matrices $a \; and \; b$ where $p = ab'$. This implies there are n-r unit roots in $p \; Y$. The VAR model can then be expressed in error correction form. That is

$$(14) \qquad \Delta Y_t = p_0 + ab' Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \Psi \; D_t + e_t$$

The matrix b' contains the cointegrating vector(s) and the matrix a has the weighting elements for the rth cointegrating relation in each equation of the VAR. The matrix rows of $b'Y_{t-1}$ are normalized on the variable(s) of interest in the cointegrating relation(s) and interpreted as the deviation(s) from the "long-run" equilibrium condition(s). In this context, the columns of a represent the speed of adjustment coefficients from the "long-run" or equilibrium deviation in each equation. If the coefficient is zero in a particular equation, that variable is considered to be weakly exogenous and the VAR can be conditioned on that variable.

4.3 THE UNRESTRICTED MODEL AND TESTING FOR COINTEGRATION

Before conducting the cointegration tests, the appropriate lag-length for the VAR must be determined and a constant model found. The lag-length is not known a priori, so some testing of lag order must be done to ensure that the estimated residuals of the VAR are white noise, that is, they do not suffer from autocorrelation, non-normality, etc. Initially, we start with a VAR that includes 4 lags on each variable, denoted VAR(4), then we estimate a VAR with 3 lags, VAR(3), and test whether the simplification from VAR(4) to VAR(3) is statistically valid. The process is repeated sequentially down to a VAR with a single lag, VAR(1).

Table 2 reports F and related statistics for testing the validity of these simplifications. The bottom block of Table 2 reports the F statistic for testing the null hypothesis indicated by the model to the right of the arrow against the maintained hypothesis indicated by the model to the left of the arrow. The p-value or the tail probability associated with the realized value of the F statistic is also reported. None of the F statistics is significant at the 1%, 5% or even the 10% critical values, and hence the simplification to a VAR with a single lag, VAR(1), is statistically valid. This result is also supported by the Schwarz criterion (SC) and the Hannan-Quinn criterion (HQ) reported in the top block of Table 2. Both criteria are minimized when the VAR has a single lag[10]. Hence, we proceed with the analysis using the VAR(1) model.

Table 3 reports summary diagnostic tests on the residuals for the VAR with a single lag, VAR(1). The diagnostic tests consist of an F-test for the null hypothesis that there is no residual vector serial correlation; a chi-square test for the null hypothesis of joint normality of the residuals; and finally two alternative F-tests for the null hypothesis that there is no residual vector heteroskedasticity.[11] The realized values of the various test statistics and the associated tail probabilities (p-values) are given in columns three and four respectively. Statistically, the VAR(1) appears well-specified, with no rejections of

[10] For an excellent reference on the various information criteria, see Judge et. al. (1988)
[11] For references on the test statistics, see Doornik and Hendry (2001).

the null hypothesis from the various test statistics. That is, the VAR residuals appear normal, homoskedastic, and serially uncorrelated.[12]

Another important aspect of diagnostic checking is testing for model constancy/stability. To accomplish that, recursive estimation techniques are employed. The basic idea behind recursive estimation is to fit the VAR to an initial sample of M-1 observations, and then fit the VAR to samples of M, M+1, …, up to T observations, where T is the total sample size.

Figure 7 shows the results from recursively estimating the VAR. Specifically, Figure 7 shows the 1-step ahead residuals for each equation of the VAR bordered by plus or minus twice their standard errors. Residuals lying outside the standard error bands are suggestive of outliers and /or model non-constancy. The estimated recursive residuals from the VAR provide evidence that the model is fairly stable. This result is confirmed by the plots of the 1-step ahead Chow tests (denoted 1up) and Break-point Chow tests (denoted Ndn), which are shown for each equation of the VAR and for the VAR system as a whole (denoted 1 up CHOWs and Ndn CHOWs respectively). The Chow statistics are scaled so that the significant critical values become a straight line at unity. That is, if the given plot exceeds unity at any point in time, this indicates a rejection of the null hypothesis of model stability at that point. The results from the various plots strongly suggest that the VAR is stable at the 1% significance level.[13] The above analysis indicates that our VAR is empirically well behaved and hence is a suitable starting point for cointegration analysis.

The cointegration analysis proceeds in several steps: testing for the existence of cointegration, interpreting and identifying the relationship(s), inference tests on the coefficients from theory and weak exogeneity. Testing permits reduction of the unrestricted general model to a final restricted model without loss of information.

[12] The diagnostic tests mentioned in the text are vector or system tests. Diagnostic tests performed on each equation of the VAR separately yield the same results as those for the entire system.
[13] See Doornik and Hendry (2001) for more details on the various Chow tests.

Table 4 presents the initial test for cointegration and is broken up into three panels. Panel A contains results on the possible number of cointegrating relations. There are four columns for the eigen-values, null hypothesis, Trace statistic, and its' associated p-value. In the first row, the null hypothesis, r=0, is that there is at most r cointegrating vectors as opposed to the alternative that there are more than zero cointegrating vectors. This hypothesis is soundly rejected with a trace statistic of 184.03 and no measurable p-value. When the possible maximum number of cointegrating relations is one against the alternative hypothesis that there is more than one, the test statistic is 58.87 and the p-value is [0.00]. This suggests that there are at least two cointegrating vectors. We cannot reject the null hypothesis that there is at most two cointegrating relations in the third row.

Panel B presents the two cointegrating vectors normalized on (domestic) patent filings and total factor productivity respectively. We decided to interpret the two vectors as a knowledge production function and function for the determinants of total factor productivity. Panel C reports the feedback coefficients and their standard errors associated with each long-run equation for the variables of the system in first differences. The cointegrating vectors or relationships as they appear are not uniquely identified and hence the standard errors of these vectors cannot be computed. Any linear combination of the two vectors forms another stationary vector, so the estimates produced by any particular vector are not necessarily unique. Therefore to achieve identification, it is necessary to impose restrictions on the cointegrating vectors. The restrictions are motivated by economic theory and enable us to test for over-identification and obtain standard errors for the over-identified parameters.

For ease of exposition and to understand the nature of the restrictions easier, the model can be written in terms of equation 14. The error term and short-run components are omitted to focus on the long-term model. Also the trend is restricted to lie in the cointegration space.

$$(15)\quad \begin{pmatrix} \Delta dp_t \\ \Delta tfp_t \\ \Delta stpl1_t \\ \Delta s\&e_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} \begin{pmatrix} 1 & -b_{12} & -b_{13} & -b_{14} & -b_{15} \\ -b_{21} & 1 & -b_{23} & -b_{24} & -b_{25} \end{pmatrix} \begin{pmatrix} dp_{t-1} \\ tfp_{t-1} \\ stpl1_{t-1} \\ s\&e_{t-1} \\ trend \end{pmatrix}$$

The ß's and a's are those reported in Panels B and C respectively in Table 4. These are the implied unrestricted long-run (cointegrating) solutions. The implied (unrestricted) long run solution of the model is given by:

$$(16)\quad dp = b_{12}\ tfp + b_{13}\ stpl1 + b_{14}\ s\&e + b_{15}\ Trend$$

$$(17)\quad tfp = b_{21}\ dp + b_{23}\ stpl1 + b_{24}\ s\&e + b_{25}\ Trend$$

Two restrictions are required to just identify the model; any additional restrictions are over-identified and thus testable. The first restriction is on the knowledge production relation and relates the dependence of new knowledge on the stock of knowledge and the R&D scientists and engineers. There does not seem to be a reason to include a direct effect from total factor productivity in this relation from the R&D based growth theory. We impose $ß_{12}$=0. Second, the literature does not suggest that current patent applications should determine productivity. This restriction is imposed by setting $ß_{21}$=0.

The above two restrictions produce a just-identified model. We consider and test three over-identifying restrictions on the ß's and a's. The first type of test is for the specification of the cointegrating relation and the latter are tests for weak exogeneity.

First, scientists and engineers in the R&D sector are unlikely to have a direct long-run impact on total factor productivity. Again, the effect is only indirect: R&D scientists and engineers produce new knowledge. New knowledge ultimately augments the stock of knowledge and the latter has a potential impact theoretically on tfp. Thus, we exclude s&e from the second cointegrating vector, by testing $b_{24} = 0$. Statistically, the likelihood ratio test of the restriction b_{24} = 0 cannot be rejected. The test statistic is $\chi^2(1) = 1.43$ [0.23]; the degrees of freedom are in parentheses and the p-value is in square brackets.

Weak exogeneity is an important issue in model reduction. It implies that inference testing can be conducted for the parameters of interest from a conditional density rather than a joint density without loss of information. The modeling effort is simpler yet still efficient. The first hypothesis is that the cointegrating relationship for new patent flows, dp_t, does not help to explain changes in total factor productivity, Δtfp_t. The restriction is a_{21} = 0. The "discovery" of new knowledge is unlikely to explain fluctuations in productivity. Our second hypothesis is that total factor productivity does not provide information about the change in the (lagged) stock of knowledge. The timing issue aside, it suggests that stpl1 is weakly exogenous with respect to the second cointegrating vector. a_{32} = 0 means that the second cointegrating relationship does not enter the equation for $\Delta stpl1_t$. These two feedback coefficients appear numerically small in panel C of table 4. If all three over-identifying restrictions are imposed, the joint hypothesis cannot be rejected: $\chi^2(3) = 4.05$ [0.26].

4.4 THE RESTRICTED COINTEGRATION MODEL

Table 5 contains the results from the three restrictions on the just identified model. It is divided into two panels. Panel A reports the restricted (and identified) estimates for the cointegrating vectors, the ß's, together with their standard errors. Panel B reports the feedback coefficients estimates, a's, and their standard errors.

In panel A, the first cointegrating vector is interpreted as a long run knowledge or "idea" production function. The implied long run or cointegrating relationship is given by:

$$(18) \quad dp = 1.436^{**}\ stpl1 + 0.208^{**}\ s\&e - 0.023^{**}\ Trend$$

The coefficient of the lagged stock of knowledge, $stpl1$, is highly significant and indicates the presence of positive spillovers of knowledge or a "standing on shoulders effect". The sign of the coefficient is consistent with the R&D-based growth models of Romer (1990) and Jones (1995b). However, its magnitude is significantly greater than unity indicating a stronger degree of spillovers than the theoretical models.[14] This result will be discussed further shortly once the second cointegrating vector has been examined.

The productivity of researchers, $s\&e$, increases with the stock of cumulated knowledge discovered by others in the past. The coefficient of $s\&e$ is positive, highly significant and less than unity. Our estimate of 0.21 is within the range 0.1 to 0.6 that Kortum (1993) finds in the micro literature on patents and R&D effort. It supports Jones (1995b) argument for decreasing returns due to duplicative research. Duplication by itself is not wasteful. Replication is an essential exercise in science and a component in learning by doing.

The negative coefficient for the time trend may seem at first surprising. However, it reflects the fact that R&D scientists and engineers grew at a much higher average annual rate than domestic patent applications over the period 1948-1997; the growth rate of the latter is 1.7% and the former is 4.3%. Their difference roughly matches the coefficient of the trend term (-2.3%). The negative coefficient of the trend term is also consistent with findings by Griliches (1990) based on micro data of patents and R&D. He interprets the negative trend as capturing a decrease in the propensity to patent inventions due to the rising cost of dealing with the patent system.

Now consider the feedback coefficients for the first cointegrating vector (the knowledge production function) in panel B of Table 5. They are all significant from zero, this means that dp, stpl1, and s&e are *not* weakly exogenous with respect to the parameters of the knowledge production function. That is, in the face of any deviation from long run equilibrium, dp, stpl1, and s&e jointly respond and move the system back to equilibrium.

---

[14] In fact the likelihood ratio statistic rejects the null hypothesis that that coefficient is unity: $\chi^2(1)= 29.12$ [0.00].

This finding supports our system approach to estimating the knowledge production function. If a single equation approach had been adopted instead, we would have invalidly conditioned on stpl1, and s&e; the result would have been biased and inconsistent estimates of the knowledge production function.

The feedback coefficient for the Δdp equation (i.e. a_{11} in equation 15) is –0.45 and significant from zero, suggesting stability of the error correction mechanism. The coefficient implies that a positive deviation of dp from its long run path (given by equation 18) this period is not permanent leading to explosive growth. The growth in (domestic) patent filings declines next period.

The feedback coefficient for the Δstpl1 equation (i.e. a_{31} in equation 15) is positive and significant from zero. The positive sign suggests that if dp is above its long run equilibrium path, then this has a positive effect on the growth of the stock of knowledge next period.

Finally, the feedback coefficient for the Δs&e equation (i.e. a_{41} in equation 15) is negative and significant from zero. The negative sign of the coefficient makes sense; it implies that if dp is above its long run equilibrium path this period, then the growth of the R&D scientists and engineers slows down in the next period to correct for the disequilibrium. This consistent with theory that that on a balanced growth path the proportion of the labor force devoted to knowledge production should remain constant.

Next, consider the second cointegrating vector in panel A of Table 5. The implied long run cointegrating relationship is given by:

$$(19) \quad tfp = 0.108 \ stpl1 + 0.009^{**} \ Trend$$

That is, total factor productivity depends positively on the lagged stock of knowledge and a time trend. The positive coefficient for the trend term is highly significant and implies a trend growth of total factor productivity of about 1%, roughly matching its average annual growth rate over the sample 1948-1997. The coefficient for the lagged stock of knowledge is positive but small. It implies that doubling the stock of knowledge will increase total factor productivity by only about 10% in the long run. The coefficient actually matches that found by Porter and Stern (2000) for aggregate data on OECD countries and falls within the range of estimates of 0.06 to 0.2 found by Griliches (1990) in the micro productivity literature.

However, the coefficient is not significant. The likelihood ratio statistic under the null hypothesis that the coefficient is zero yields: $\chi^2(1) = 0.62$ [0.43]. One possible interpretation from Porter and Stern is that realizing the full benefits from new knowledge and new technologies depends critically on the diffusion of these technologies into the productive sector of the economy. They provide some evidence that the rate of diffusion of new technologies has been slow and incomplete in the OECD countries over the past 20 years. That may explain the weak relationship between the stock of knowledge and TFP. In fact, this idea is supported by or appears to be evident with the US data depicted in Figure 8. Since the mid 1980s, it does not look like the benefits of knowledge have been realized into measured productivity growth. Kortum identifies the trend shift in patenting behavior. We suspect that the diffusion process of this trend in knowledge growth did not have its full effect on total factor productivity until the late 1990s just beyond our sample.

The cointegrating relation for total factor productivity enters into the Δtfp_t equation with the appropriate sign and is significant. The feedback effect is rather quick. Changes in scientists and engineers, $?s\&e_t$, are not weakly exogenous as well. If tfp is above the long-run level, it reduces s&e growth over several years. There is less “demand” for knowledge producing workers. This has a negative effect on changes in patents, $?dp_t$, and is seen in the marginally significant feedback coefficient as well. The stock of knowledge is weakly exogenous with respect to the tfp cointegrating relation. We are not as confident of the interpretation of the feedback coefficients from the total factor productivity relation, because the relationships are perhaps more complex and there is not as much direction from economic theory.

There are two important results mentioned above that we want to focus on in the remainder of this section. The first is the estimated long run impact of the lagged stock of total patents on domestic patents. This is the coefficient of stpl1 in equation (18), which governs the magnitude of knowledge spillovers. The coefficient is 1.4, significantly greater than unity. The second is the estimated long run impact of the lagged stock of total patents on total factor productivity. This is the coefficient of stpl1 in equation (19), which is 0.108.

It is hard to compare the knowledge spillovers coefficient to the micro literature on patents and R&D, because the knowledge production function from that literature does not derive from the R&D-based models of growth and hence does not include the stock of existing knowledge as a variable explaining new knowledge. Also, simple aggregation from the micro level may not capture the potential externalities across sectors.

However, Porter and Stern (2000), using aggregate panel data on OECD countries, do estimate a knowledge production function, where domestic patents in each country do depend on the existing stock of knowledge of that country. However, they use as a proxy for the stock of knowledge in a particular country the *stock of domestic patents* for that country rather than the stock of total (domestic and foreign) patents, which we use for the US economy. Their results indicate a spillover coefficient of unity supporting the Romer (1990) model.

As a check of robustness of our model, we do re-estimate a knowledge production function but now we replace the stock of total (domestic and foreign) patents by the stock of domestic patents. To explain, let sdpl1 denotes the (lagged) *stock* of domestic patents in the US economy. Starting from an unrestricted VAR that includes dp, tfp, sdpl1, and s&e, we implement the Johansen cointegration methodology and again estimate 2 cointegrating vectors. The steps of this exercise are exactly the same as those for the main model discussed in this paper. (The results are reported in table 6). Here, we focus on what is relevant for the present discussion. We find that the estimated long run impact of the (lagged) *stock* of domestic patents on the flow of new domestic patents (dp) to be unity and significant, but still the long run impact of the (lagged) *stock* of domestic patents on total factor productivity is small and insignificant (the coefficient is 0.1 with a standard error of 0.1).

The result that the impact of the (lagged) *stock* of domestic patents on the flow of new domestic patents is unity is consistent with Porter and Stern (2000) and Romer (1990). This result actually supports the coefficient of 1.4 on stpl1 in equation (18): when the stock of total (domestic+ foreign) patents is used to measure the stock of knowledge, researchers have a larger pool of knowledge to draw upon, and hence the spillover effect is stronger than the case where the stock of domestic patents is used.

We still need to reconcile these results with the time series evidence presented by Jones (1995a). Recall first that the Romer (1990) model assumes a spillover parameter of unity with the implication that the growth rate of the stock of knowledge is proportional to the R&D scientists and engineers. Jones (1995a) directly measured the stock of knowledge using total factor productivity and rejected Romer's (1990) assumption of a spillover parameter of unity on the basis of the observed weak relationship between TFP growth and the number of R&D scientists and engineers: the number of R&D scientists and engineers has been trending strongly upward over the postwar period with no apparent benefit in terms of faster TFP growth. Jones argued that this weak relationship is therefore an indication of weak knowledge spillovers.

The empirical results presented in this paper suggest that once patents are used to measure knowledge and hence the knowledge production function is directly estimated, the knowledge spillover parameter is unity or larger. That is, R&D scientists and engineers appear to have greatly benefited from the knowledge and ideas discovered by prior research. Therefore, the observed weak relationship between TFP growth and R&D scientists and engineers documented by Jones is not necessarily inconsistent with the presence of large knowledge spillovers once knowledge is measured using patent statistics.

The results of this paper point to an alternative explanation of this weak relationship: the knowledge that R&D scientists and engineers produce seems to have had only a limited impact on measured total factor productivity. Productivity increases are the result of the successful application and the embodiment of knowledge. In a simple OLS case, we could attribute some of the weak result(s) to measurement problems. Errors in the dependent variable raise the standard error reducing the precision of the estimates. Increases in knowledge are not homogeneous or necessarily applicable in an economic sense. This results in attenuation bias for the estimates reducing their size (and consistency). An errors in variables problem may partially explain the weak relationship between the stock of knowledge and TFP. However, the dynamic approach from the Johansen approach may minimize the effect as the lagged variables act like instruments. Our empirical work contributes to understanding and reconciling some of the spillover effects and issues raised by Jones (1995a).

5. RECONCILIATION AND CONCLUSION

This paper began with a theoretical presentation of R&D-based models of economic growth. The paper showed that this class of models derives long run growth through the accumulation of knowledge/technology. Knowledge, in turn, is a variable whose evolution is modeled endogenously. At any point in time, the rate of production of new knowledge depends positively on the existing stock of knowledge and the number of research workers.

The dependence of new knowledge on the existing stock captures knowledge spillovers. A large magnitude for knowledge spillovers implies that long run growth is potentially dependent on policy variables such as subsidies to R&D [Romer (1990)]. On the other hand, a small magnitude for knowledge spillovers implies that long run growth is invariant to policy [Jones (1995b)]. Hence, empirically assessing the strength of knowledge spillovers is important, at least from a policy perspective.

There is an extensive empirical micro literature that looks at different kinds of spillovers from R&D activities. These include market spillovers, network spillovers, as well as knowledge spillovers [see for example Jaffe et. al. (2000)]. In this paper, we did not consider market and network spillovers. Rather, the analysis was confined to testing the magnitude of knowledge spillovers at the macro level, as postulated by the R&D-based growth literature. The existing empirical macro research on the magnitude of knowledge spillovers seems to suggest that they are small.

In this paper, we focused on estimating the parameters of the knowledge production function. This allowed us to directly assess the magnitude of knowledge spillovers. To achieve this goal, we exploited historical time series of patent applications to construct knowledge flows and stocks. Cointegration techniques were used for analyzing and estimating models that involve non-stationary data and the specified long-run relationships.

We found evidence for two potential long-run relationships. The first is interpreted as a knowledge production function; the second is a relationship that captures a positive dependence of total factor productivity on the stock of knowledge. The results indicate that knowledge spillovers are large. The estimated long run elasticity of new knowledge with respect to the existing stock of knowledge is unity when the domestic stock of patents is used a proxy for the knowledge stock and greater than unity with a global measure of knowledge based on both domestic and foreign stocks of patents. However, the estimated long run elasticity of total factor productivity with respect to the stock of knowledge is positive, small and imprecise.

Our empirical work contributes to understanding and reconciling some of the spillover effects and issues raised by Jones (1995a).These results seem to suggest that while research workers benefit greatly from "standing on the shoulders" of prior researchers, the knowledge that they produce seems to have complex and slowly diffusing impacts on total factor productivity. Therefore, the results suggest a new interpretation of the empirical evidence. The observed weak relationship between TFP growth and the number of research workers he found does not necessarily indicate that knowledge spillovers are small. Rather, it is possibly due to the small impact knowledge has on TFP. In addition, we might attribute some of the weak result(s) to measurement problems in both variables. The application and embodiment of knowledge into productivity is complex and diffuses slowly (even in the long-run). Further work can examine the transmission mechanism(s) between the knowledge production function with productivity.

REFERENCES

Aghion, Philippe, and Peter Howitt (1992). " A Model of Growth through Creative Destruction." *Econometrica* 60, 323-51.

Arora, A., and A. Cambardella (1994). " The Changing Technology of Technological Change: General and Abstract Knowledge and the Division of Innovative Labor." *Research Policy*, 23, 523-532.

Coe, David T., and Elhanan Helpman (1995). " International R&D Spillovers." *European Economic Review* 39, 859-887.

Davidson, J.E.H., David F. Hendry, F. Srba, and S. Yao (1978), "Econometric Modeling of the Aggregate Time Series Relationship Between Consumer's Expenditure and Income in the United Kingdom," *Econometric Journal*, 88, 661-692.

Doornik, Jurgen A., David F. Hendry (2001). Modelling Dynamic Systems Using PcGive 10, Vol. 2. London: Timberlake Consultants Press.

Doornik, Jurgen A., David F. Hendry, and Bent Nielson (1998). " Inference in Cointegrating Models: UK M1 Revisited." *Journal of Economic Surveys* Vol.12, No.5, 533-572.

Enders, W. (1995). Applied Econometric Time Series. John Wiley & Sons, Inc., New York.

Greenwood, J. and M. Yorukoglu (1997). "1974." *Carnegie-Rochester Conference Series on Public Policy* 46, 49-95.

Griliches, Zvi (1989). " Patents: Recent Trends and Puzzles." *Brookings papers on Economic Activity, Microeconomics*, 291-330.
Griliches, Zvi (1990). " Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature* 18, 4, 1661-1707.

Grossman, Gene, and Elhanan Helpman (1991a). Innovation and Growth in the Global Economy. Cambridge, Massachusetts and London: MIT Press.

Grossman, Gene, and Elhanan Helpman (1991b). " Quality Ladders in the Theory of Growth." *Review of economic Studies* LVIII, 43-61.

Hausman, J.A., B.H. Hall, and Z. Griliches (1984). " Econometric models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 52 (4), 909-38.

Hendry, David F., 1986, " Econometric Modelling with Cointegrated Variables: An Overview," Oxford Bulletin of Economics and Statistics, 48, 3, pp 201-12.

Jaffe, Adam, M. Trajtenberg, and M. Fogarty (2000). “ Knowledge Spillovers and Patent Citations: Evidence From a Survey of Inventors.” *The American Economic Review Papers and Proceedings*, May.

Jaffe, Adam, Rebecca Henderson, and Manual Trajtenberg (1993). “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *The Quarterly Journal of Economics,* 108 (3), 577-598.

Johansen, S. (1988). “ Statistical Analysis of Cointegrating Vectors.” *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S. (1991). “ Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models.” *Econometrica* 59, 1551-80.

Jones, Charles (1995a). “ Time Series Tests of Endogenous Growth Models.” *Quarterly Journal of Economics* 110, (May), 494-525.

Jones, Charles (1995b). “ R&D-Based Models of Economic Growth.” *Journal of Political Economy* 103, (August), 759-84.

Jones, Charles I. (2002). “ Sources of U.S. Economic Growth in a World of Ideas.” *American Economic Review* Vol. 92, No. 1, 220-239.

Joutz, Frederick L., and Thomas A.Gardner (1996). “ Economic Growth, Energy Prices and Technological Innovation.” *Southern Economic Journa*l Vol. 62, No. 3, 653-666.

Judge, G., R. Hill, W. Griffiths, H. Lutkepohl, and T. Chao Lee (1988). Introduction to the Theory and Practice of Econometrics. Second Edition. John Wiley & Sons, Inc., New York.

Kortum, Samuel S. (1993). “ Equilibrium R&D and the Patent-R&D Ratio: U.S. Evidence.” *American Economic Review* Vol. 83, 2, 450-457.

Kortum, Samuel S. (1997). “ Research, Patenting, and Technological Change.” *Econometrica* Vol. 65, No. 6, 1389-1419.

Kortum, Samuel S., and Josh Lerner (1998). “ Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?” *Carnegie-Rochester Conference Series on Public Policy* 48, 247-304.

Machlup, Fritz (1962). The Production and Distribution of Knowledge in the United States. Princeton, New Jersey: Princeton University Press.

Nelson, C.R. and C.I. Plosser, 1982, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications,” Journal of Monetary Economics, Vol. 10, No. 1, February, pp. 139-162.

NSF (1955). *Scientific Research Personnel Resources.* National Science Foundation, Washington, D.C.

NSF (1961). *Reviews of Data on Research and development.* National Science Foundation, Washington, D.C.

NSF (1962). *Reviews of Data on Research and development.* National Science Foundation, Washington, D.C.

NSF (1993). *Science and Engineering Indicators- 1993.* National Science Foundation, Washington, D.C.

NSF (1998). *National Patterns of R&D Resources: 1998.* National Science Foundation, Washington, D.C.

Perron, Pierre (1989). " The Great Crash, The Oil Price Shock, and The Unit Root Hypothesis." *Econometrica* Vol. 57, No. 6, 1361-1401.

Porter, Michael E., and Scott Stern (2000). " Measuring the "Ideas" Production Function: Evidence from International Patent Output." *National Bureau of Economic Research Working Paper*, No.7891.

Romer, Paul M. (1990). "Endogenous Technological Change." *Journal of Political Economy* 98, 5 (October): pt. II, 71-102.

Thompson, Peter, and Melanie Fox Kean (2004) "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review*, 21 (1), 172-187.

Figure 1a: Annual growth rates for total factor productivity (differenced log).

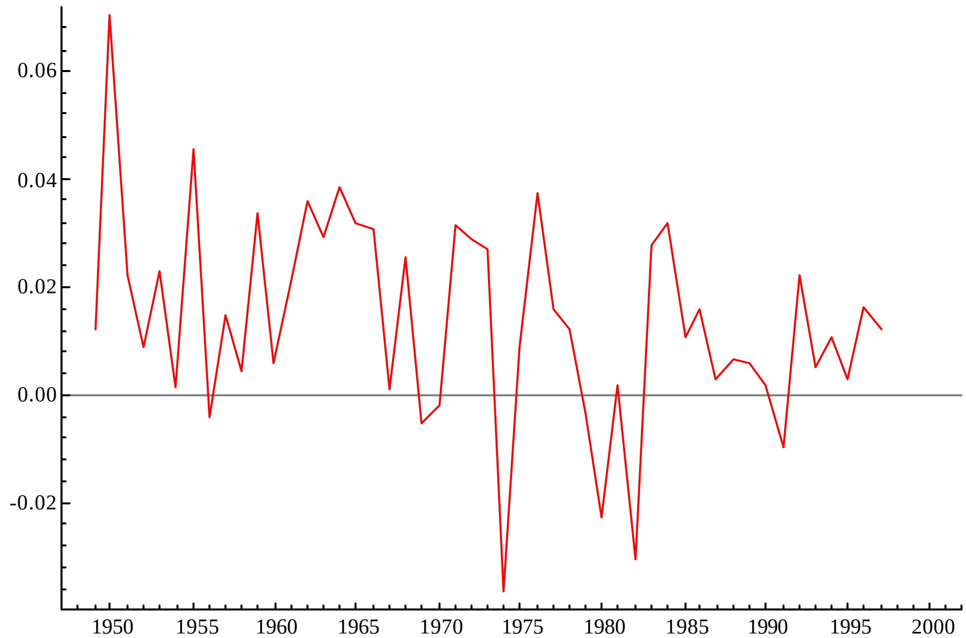


Figure 1b:

The number of scientists and engineers engaged in R&D in the US (in thousands).

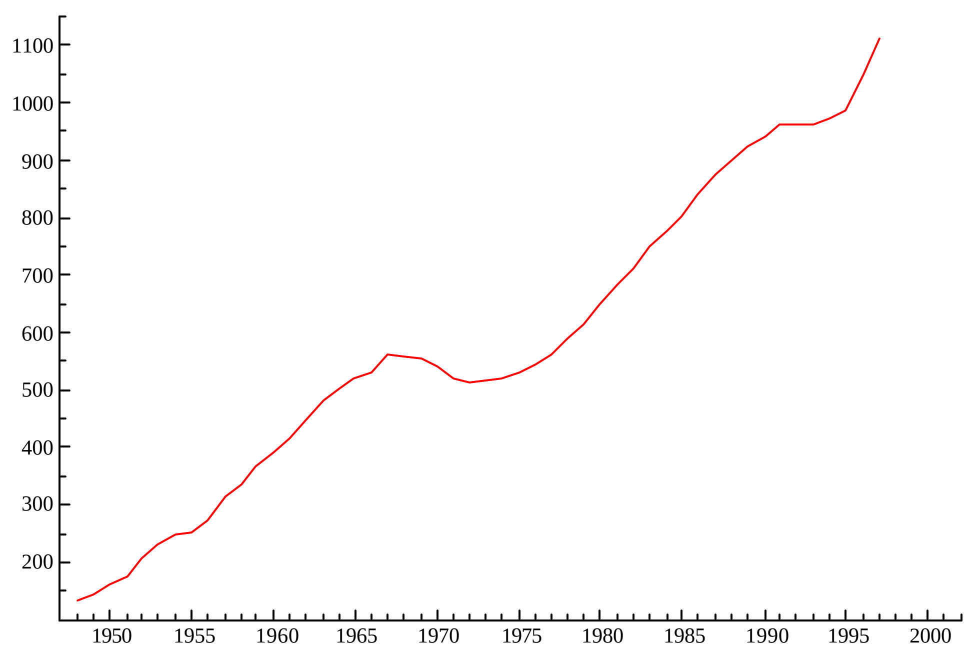


Figure 2: The log of domestic patent applications (dp).

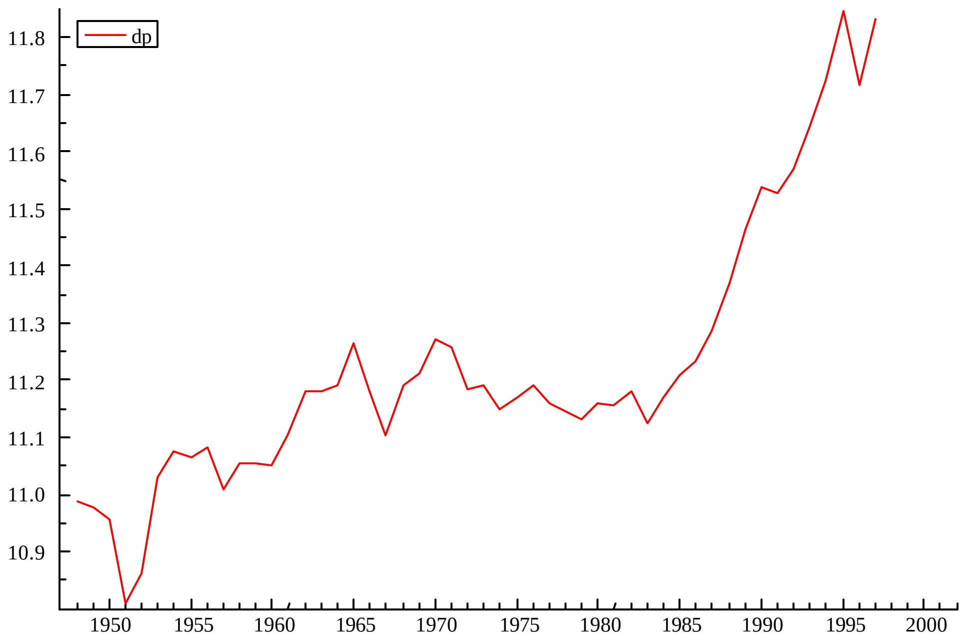


Figure 3: The log of the stock of total patent applications (stp).

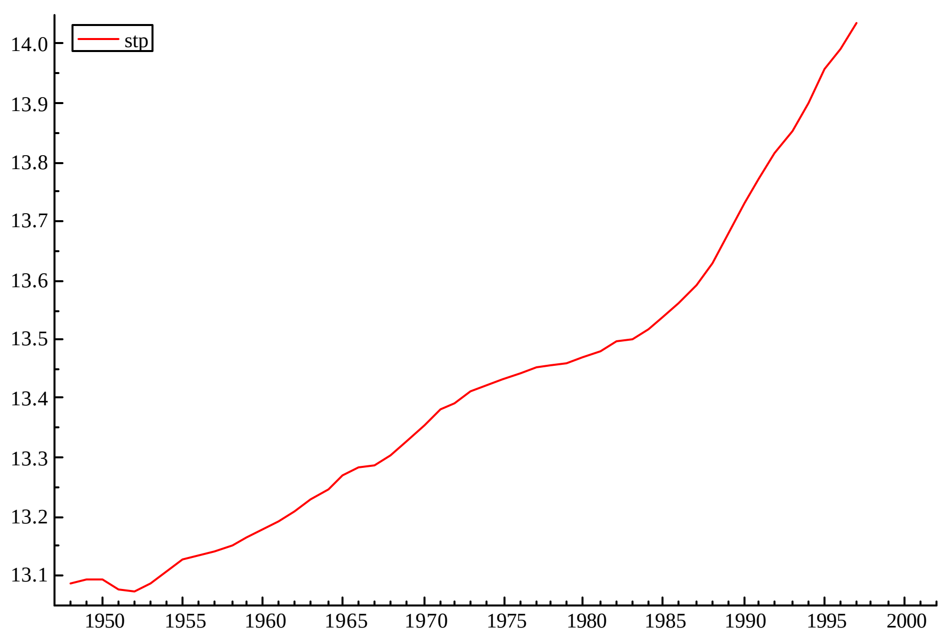


Figure 4: The log of the number of scientists and engineers engaged in R&D (s&e).

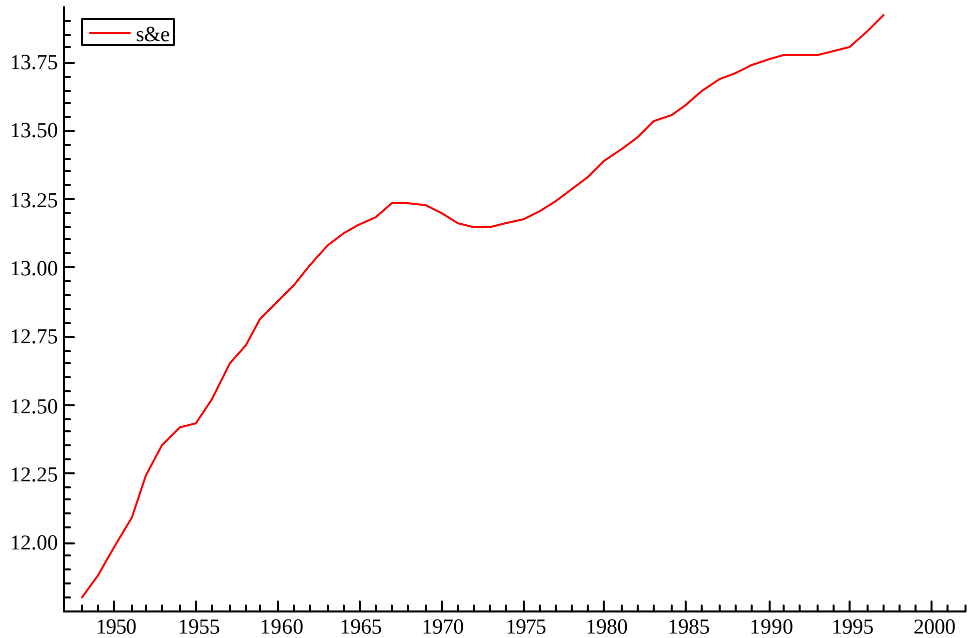


Figure 5: The log of total factor productivity (tfp).

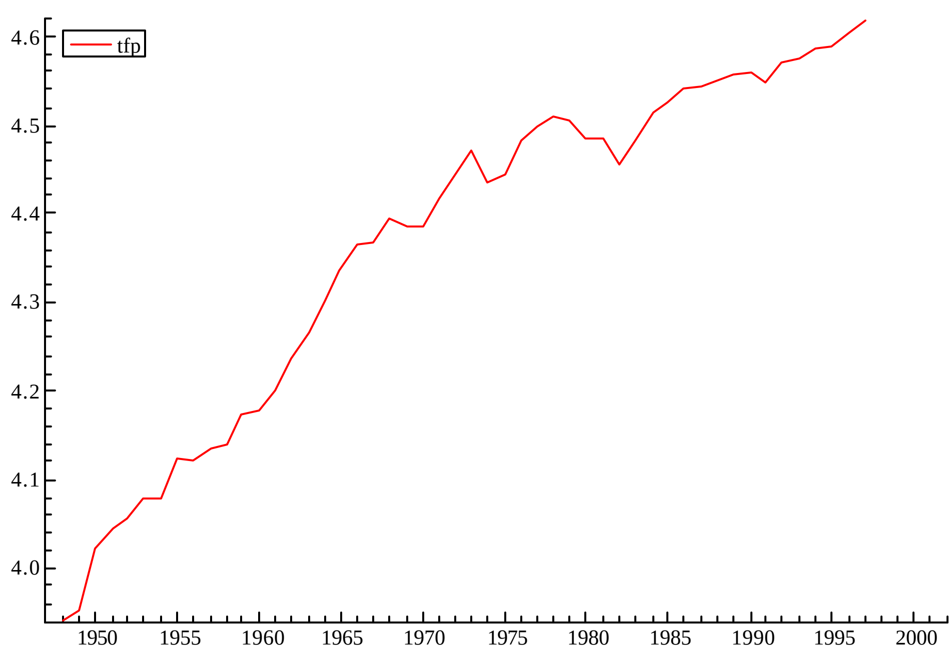


Figure 6:
The first difference of the log of the stock of total patent applications (?stp).

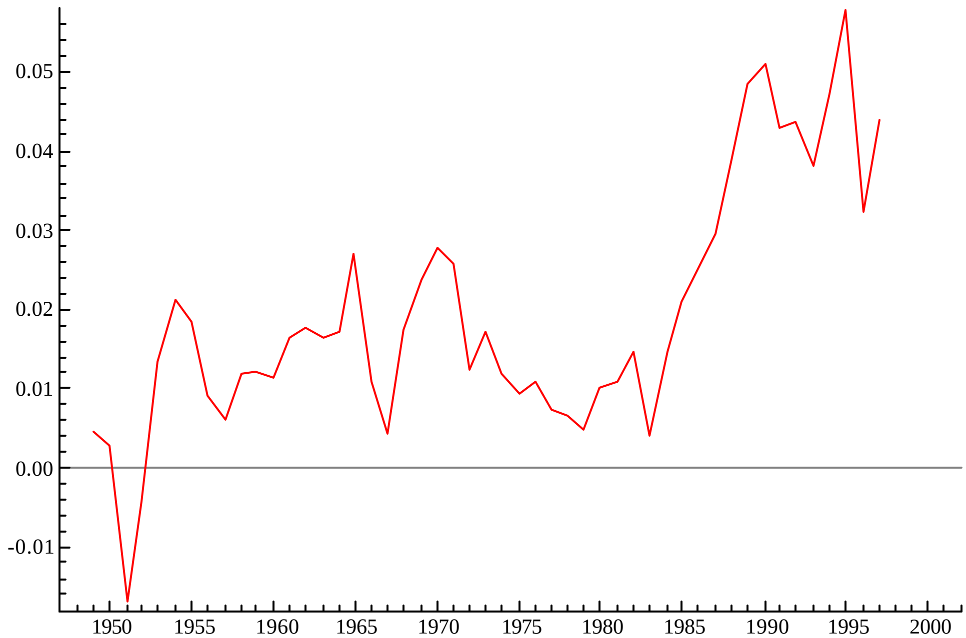
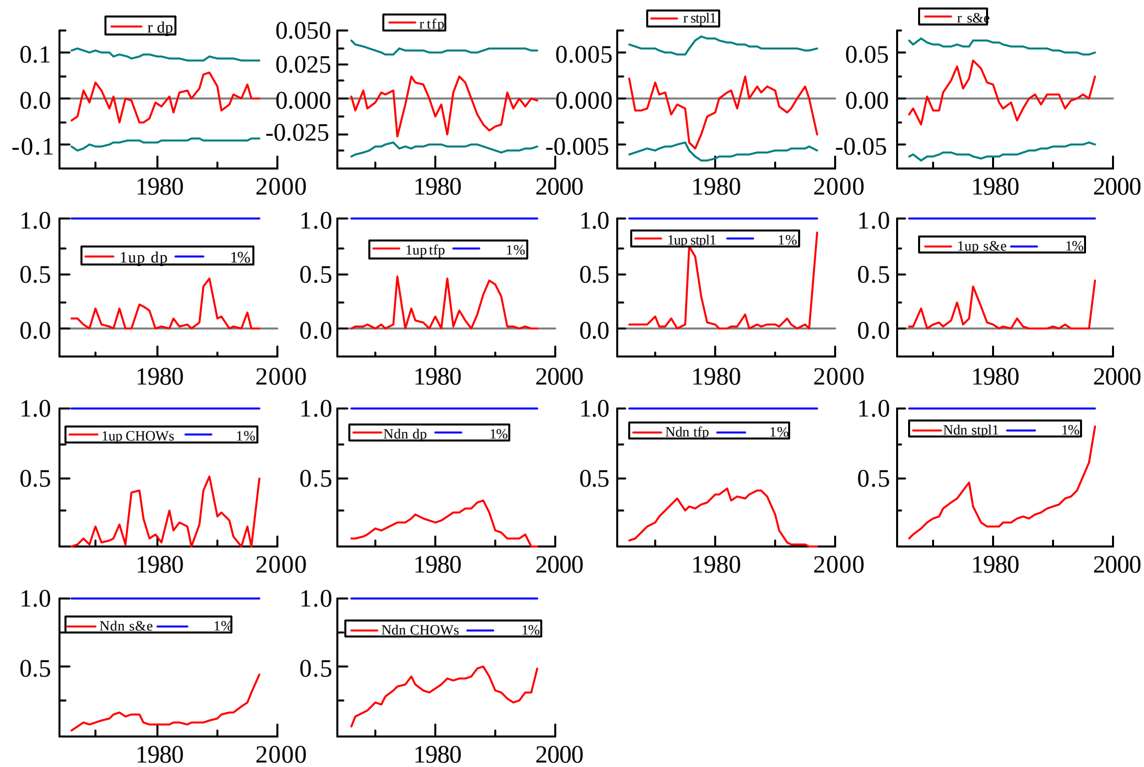


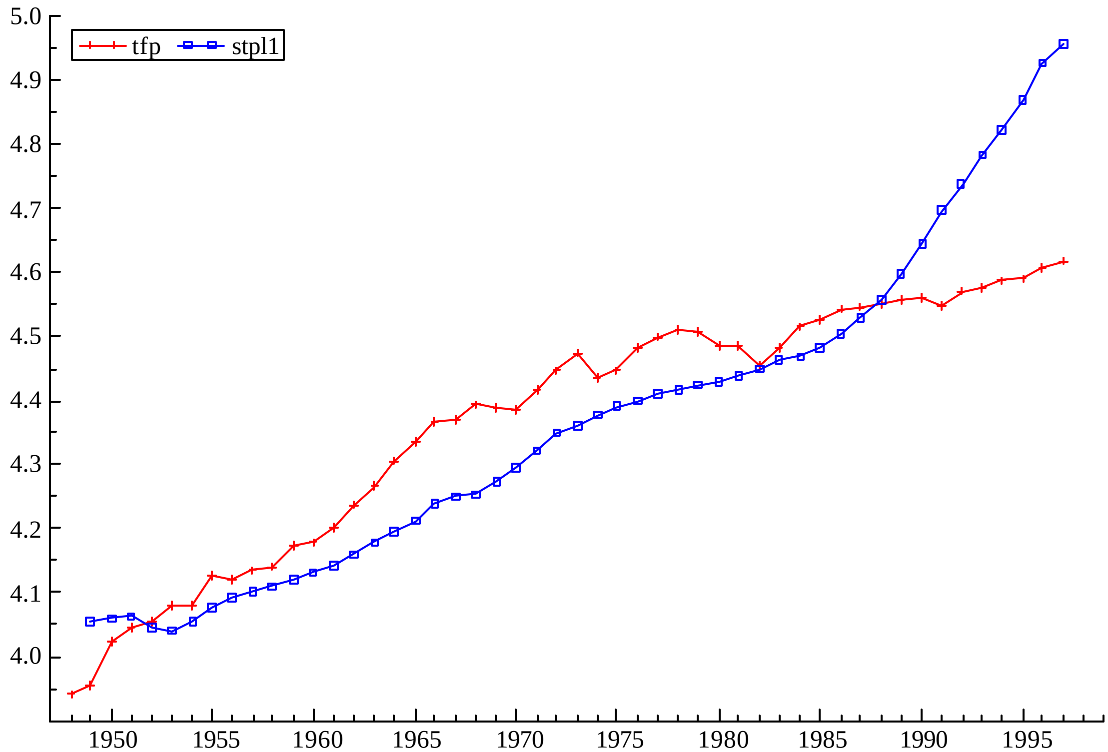
Figure 7: Model stability tests: Recursive residuals and Chow tests.



Notes:

(1) The VAR includes a single lag on each variable (dp, tfp, stpl1, s&e), a constant, trend, and three dummy variables: Stepdum86, Impluse9495, and Impulse96. The estimation sample is: 1953 to 1997.

Figure 8: Total factor productivity and the (lagged) stock of total patents (in logs).



Notes:

The series are adjusted by their means.

Table 1

Augmented Dickey-Fuller Test Results for Levels and Differences, 1953-1997

Variable	Model with a Constant		Model with a Constant and Trend	
	Lag-length	t-ADF	Lag-length	t-ADF
dp	0	0.22	0	-0.78
tfp	0	-2.25	0	-1.43
s&e	1	0.89	1	-2.05
stp	1	1.7	1	0.44
sdp	1	0.77	1	-0.44
? dp	0	-6.51**	0	-6.52**
? tfp	0	-6.71**	0	-6.59**
? s&e	0	-3.75**	0	-3.36*
? stp	0	-2.03	0	-2.64
? stp (Perron)	1	-2.12	1	-2.08
? sdp	0	-2.04	0	-2.32
? sdp (Perron)	1	-2.00	1	-1.93

Notes:

(1) For a given variable x, the augmented Dickey-Fuller equation with a constant term included has the following form:

$$\Delta x_t = p \ x_{t-1} + \sum_{i=1}^{p} q_i \ \Delta x_{t-i} + a + e_t$$

where e_t is a white noise disturbance. For each variable, the table reports the number of lags on the dependent variable, p, chosen using the Akaike information criterion. The table also reports the augmented Dickey-Fuller statistic, t-ADF, which is the t-ratio on π from the above regression. The statistic tests the null hypothesis of a unit root in x, i.e. π=0, against the alternative of stationarity. Critical values at the 5% &1% significance levels respectively are –2.927 and –3.581.

(2) The augmented Dickey-Fuller equation with a constant and trend included has exactly the form in (1) with an additional trend term included as a right-hand- side variable. Again, for each variable, the table reports the number of lags on the dependent variable and the augmented Dickey-Fuller statistic. Critical values at the 5% and 1% significance levels are respectively –2.927 and –3.581.

(3) The symbols * and ** denote rejection of the null hypothesis at the 5% and 1% critical values respectively.

(4) The Perron adjusted results report the test for stationarity with a structural shift in the mean with the break point at 1985, approximately 80% from the starting observation. The critical values tabulated by Perron (1989) are -3.82 and -4.38 at the 5% and 1% significance levels respectively.

Table 2

F and Related Statistics for Testing Model Reduction

Model	Log Likelihood	Schwartz criterion	Hannan-Quinn criterion
VAR(4)	590.0	-19.12	-21.23
VAR(3)	571.2	-19.64	-21.35
VAR(2)	555.2	-20.28	-21.59
VAR(1)	539.4	-20.93	-21.84

Model Reduction	Statistic	Value	p-Value
VAR(4) --> VAR(3):	F(16,64) =	1.28	[0.241]
VAR(3) --> VAR(2):	F(16,77) =	1.26	[0.246]
VAR(4) --> VAR(2):	F(32,79) =	1.29	[0.184]
VAR(2) --> VAR(1):	F(16,89) =	1.45	[0.139]
VAR(3) --> VAR(1):	F(32,93) =	1.37	[0.124]
VAR(4) --> VAR(1):	F(48,82) =	1.37	[0.104]

Notes:

(1) The Var(.) includes the four main variables: dp, tfp, stpl1, s&e. It also includes a constant, trend, and the three dummy variables: Stepdum86, Impluse9495, Impulse96. The estimation sample is: 1953 to 1997.

(2) The bottom block reports the F statistic testing the null hypothesis indicated by the model to the right of the arrow against the maintained hypothesis given by the model to the left of the arrow. The tail probabilities associated with the values of the F statistic are reported in square brackets.

Table 3

Summary Diagnostic Test Statistics for VAR(1) Residuals

Test	**Statistic**	**Value**	**p-Value**
Vector AR 1-2 test:	F(32,93)	1.04	[0.432]
Vector Normality test:	Chi^2(8)	6.53	[0.588]
Vector hetero test:	F(100,126)	0.71	[0.962]
Vector hetero-X test:	F(200,76)	0.55	[1.000]

Notes:

(1) The VAR includes a single lag on each variable (dp, tfp, stpl1, s&e), a constant, trend, and three dummy variables: Stepdum86, Impluse9495, and Impulse96. The estimation sample is: 1953 to 1997.

Table 4

Cointegration Analysis of the Data

A) Johansen's Cointegration Test

Eigenvalues	**Null Hypothesis**	**Trace Statistic**	**p-Value**
0.937	r = 0	183.03**	[0.000]
0.687	r ≤ 1	58.87 **	[0.000]
0.120	r ≤ 2	6.65	[0.993]
0.019	r ≤ 3	0.88	[0.997]

(B) Estimated Cointegrating Vectors b¢

Vector	**dp**	**tfp**	**stpl1**	**s&e**	**Trend**
1	1	-0.255	-1.415	-0.195	0.025
2	0.170	1	-0.418	-0.106	-0.002

(C) Feedback Coefficients a **and Their Standard Errors SE(**a **)**

	a		**SE(**a **)**	
	1	**2**	**1**	**2**
dp	-0.398	-0.342	0.120	0.121
tfp	0.062	-0.109	0.050	0.050
stpl1	0.132	0.022	0.008	0.008
s&e	-0.162	-0.402	0.071	0.072

Notes:

(1) The VAR includes a single lag on each variable (dp, tfp, stpl1, s&e), a constant, trend, and three dummy variables: Stepdum86, Impluse9495, and Impulse96. The estimation sample is: 1953 to 1997.

Table 5

Restricted Cointegration Analysis of the Data

(A) Cointegrating Vectors b¢ **and Their Standard Errors (in parentheses)**

Vector	**dp**	**tfp**	**stpl1**	**s&e**	**Trend**
1	1	0.0	-1.436**	-0.208**	0.023**
			(0.0619)	(0.020)	(0.001)
2	0.0	1	-0.108	0.0	-0.009**
			(0.141)		(0.002)

(B) Feedback Coefficients a **and Their Standard Errors SE(**a **)**

	a		**SE(**a **)**	
	1	**2**	**1**	**2**
dp	-0.449**	-0.184	0.121	0.111
tfp	-----	-0.108*	-----	0.050
stpl1	0.135**	-----	0.008	-----
s&e	-0.190**	-0.405**	0.068	0.053

Notes:

(1) The VAR includes a single lag on each variable (dp, tfp, stpl1, s&e), a constant, trend, and three dummy variables: Stepdum86, Impluse9495, and Impulse96. The estimation sample is: 1953 to 1997.

(2) * and ** indicate the rejection (at the 5% and 1% critical values) of the null hypothesis that a particular coefficient is zero. These tests are based on the likelihood ratio statistic, which is distributed under the null hypothesis as χ^2 with 1 degree of freedom.

Table 6

Restricted Cointegration Analysis of the Data

(A) Cointegrating Vectors b¢ **and Their Standard Errors (in parentheses)**

Vector	**dp**	**tfp**	**sdpl1**	**s&e**	**Trend**
1	1	0.0	-1.000**	-0.011*	0.00046*
			(0.009)	(0.004)	(0.00016)
2	0.0	1	-0.134	0.0	-0.0088**
			(0.127)		(0.0013)

(B) Feedback Coefficients a **and Their Standard Errors SE(**a **)**

	a		**SE(**a **)**	
	1	**2**	**1**	**2**
dp	-0.341**	-0.356**	0.109	0.116
tfp	-----	-0.102*	-----	0.048
sdpl1	0.162**	-----	0.001	-----
s&e	-0.166**	-0.413**	0.060	0.069

Notes:

(1) The VAR includes a single lag on each variable (dp, tfp, sdpl1, s&e), a constant, trend, and three dummy variables: Stepdum86, Impluse9495, and Impulse96. The estimation sample is: 1953 to 1997.

(2) * and ** indicate the rejection (at the 5% and 1% critical values) of the null hypothesis that a particular coefficient is zero. These tests are based on the likelihood ratio statistic, which is distributed under the null hypothesis as χ^2 with 1 degree of freedom.

APPENDIX A: DATA CONSTRUCTION AND SOURCES

Stock of total (domestic and foreign) patent applications:

The US Patent and Trademark Office provides information on the number of patent applications filed from 1840 to present. These include patents for invention, designs, and plants. This data is available on-line from the US Patent and Trademark Office web site, at:

http://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm

Patents for invention, designs, and plants summed up to get the total number of patent applications. The total number of patent applications is used to construct a patent stock measure following the methodology proposed by Joutz and Gardner (1996) whereby:

(A1) $Stock_t = (\text{total number of patent applications})_t + Stock_{t-1}\,(1-d)$

where d denotes the depreciation rate. To make the above formula operational, an initial stock needs to be estimated. We calculate the initial stock of patents using two alternative methodologies. In the first, we follow the procedure suggested by Coe and Helpman (1995) and calculate the initial stock as:

$$Stock_0 = (\text{number of patent applications})_0 / (g+d)$$

where g is the average annual growth rate of the number of patent applications over the period for which data is available (1840-1999). The (number of patent applications)$_0$ denotes the number of patent applications in the first year for which data is available (1840), and $Stock_0$ is the stock for that year (1840). Alternatively, we simply assume that the initial stock in 1840 is equal to the number of patent applications in 1840. That is, we assume that prior to 1840 there was no “Knowledge”.

With an assumed depreciation rate, usually 15 % [see Griliches (1989, 1990)], and an estimate for the initial stock in 1840, equation (A1) is used to construct the subsequent patent stocks. It turns out that both procedures for calculating the initial stock yield virtually identical patent stock series over the postwar period, and hence, it does not matter which one is used.

Also, stocks were constructed using alternative depreciation rates typically used in the literature (0, 5, 10%). The results are consistent and robust across the range. [These findings are available upon request.]

Domestic patent applications:
From 1940 to 1999, the data is obtained from the following sources:

- Tabulations of the US Patent and Trademark Office available online at http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.pdf

- Unpublished memorandum of P.J. Federico, the US Patent and Trademark Office, January 18, 1961. This was obtained from Jim Hirabayashi at the Technology Assessment and Forecast Branch, the US Patent and Trademark Office.

- *Journal of the Patent Office Society,* Vol 44 (No. 2), February 1964, page 168.

- Commissioner of Patents and Trademarks, Annual Report, 1966, page 26, the US Patent and Trademark Office.

From 1840 to 1939, data on domestic patent applications is not available. We follow Kortum and Lerner (1998) and proxy for the number of domestic patent applications by multiplying the total number of patent applications by the fraction of total patent grants issued to US inventors.

Stock of domestic patent applications:
We cumulate the number of domestic patent applications into a stock measure using the perpetual inventory method with an assumed depreciation rate of 15%.

Scientists and Engineers engaged in R&D:

The data for the period 1979-1997 is obtained from the National Science Foundation, *Science and Engineering Indicators- 2000*. This source is available online at: http://www.nsf.gov/sbe/srs/seind00/start.htm. For years prior to 1979, the data is taken from Jones (2002) and Machlup (1962) who in turn obtain their data from the National Science Foundation [NSF (1993, 1962, 1961, and 1955)].

Total factor productivity (TFP):

The data is for the private business sector of the U.S. economy. It was obtained from Larry Rosenblum at the Office of Productivity and Technology, the Bureau of Labor Statistics, U.S. Department of Labor.