Outsourced R&D and GDP growth

By ANNE MARIE KNOTT *

March 9, 2015

Abstract: *In January 2011, President Obama signed into law the "America COMPETES Reauthorization Act of 2010", whose goal was to restore US competitiveness through investment in R&D. This act reflects the belief (and corresponding economic theory) that technological progress drives economic growth. However coarse comparison between R&D and GDP growth over the past forty years indicates that scientific labor has increased 2.5 times, while GDP growth has at best remained stagnant. The leading theory to explain the disconnect holds that R&D has gotten harder. I develop and test an alternative view that firms have become worse at it. Using a novel measure, RQ, I show a 65% decline in R&D productivity. Moreover I show that much of the decline stems from increased outsourcing of R&D, which is unproductive for the funding firm. I find no evidence that R&D has gotten harder—the productivity of internal R&D is essentially unchanged. This offers hope R&D productivity and growth may be fairly easily restored by bringing outsourced R&D back inhouse.*

*Knott: Washington University, One Brookings Drive, Campus Box 1133, St Louis, MO 63130 (email:knott@wustl.edu),

Acknowledgments: I gratefully acknowledge the support of the National Science Foundation under NSF Award 1246893: The Impact of R&D Practices on R&D Effectiveness. I also would like to thank researchers at the U.S. Census Bureau for their help reviewing data. This paper benefits from comments received during presentations at Washington University, Georgia Tech and the Census RDC Annual Research Conference.

DISCLAIMER: Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed

DISCLAIMER: The author has financial interest in amkANALYTICS which provides RQ data and diagnostics to firms and the investment community.

I. Introduction

In January 2011, President Obama signed into law the "America COMPETES Reauthorization Act of 2010". The goal of the act was to invest in innovation through research and development (R&D) and to thereby improve the competitiveness of the United States. This act reflects the belief since at least Solow (1957) that technological progress drives economic growth. However coarse comparison between R&D and GDP growth over the past forty years (Figure 1) indicates that scientific labor has been increasing, while GDP growth has been declining.

[Insert Figure 1 Here]

This disconnect between R&D and growth conflicts with expectations from endogenous growth theory (Romer 1990, Grossman and Helpman 1991, Aghion and Howitt 1992). Endogenous growth theory builds upon the basic insight of Solow's model (and empirics) that technological progress drives economic growth. However it departs from Solow with regard to the source of technological change. Solow treated technological change as exogenous (which would be the case if it were coming from universities or government labs), whereas endogenous growth theory treats technological change as the outcome of investment by profit maximizing agents. One of the predictions from Romer is "scale effects"—that the growth rate, g, of knowledge, A, (and under balanced growth, the growth rate for capital and final goods output as well) is defined by the total human capital employed in research, H_A, and the productivity per researcher, δ:

$$g = \delta H_A \tag{1}$$

Thus Romer expects growth to increase in the level of R&D. Figure 1 suggests that isn't happening in the US. Jones (1995), who first compared the trends in Figure 1, concludes that the "prediction of scale effects is clearly at odds with time series empirical evidence". He and others (Kortum 1997, Segerstrom 1998) propose that the disconnect between theory and the empirical record occurs because R&D has gotten harder. In an effort to reconcile theory with the data, Jones offers a model that preserves the basic structure from endogenous growth, but eliminates the scale effects prediction. In particular he introduces a "fishing out" term, θ, which causes the rate of innovation to decrease with the level of knowledge. He also introduces an "externalities" term, λ, which captures duplication in the R&D process by reducing innovation per unit of scientific labor, H_A. Adding fishing out and externalities to the basic growth model yields revised expectations for growth:

(2) $$g = \delta H_A^{\lambda-1} A^{\theta}$$

In Jones' view, Romer is a specific form of this more general functional form in which λ and θ both equal 1. If Jones' arguments are valid then the rate of growth is declining in the level of knowledge, A, as well as the amount of research labor, H_A. Thus growth converges towards zero.

Because the Jones result that growth is independent of R&D runs counter to the intent of endogenous growth theory, Peretto (1998) and others (Young 1998, Aghion and Howitt 1998, and Dinopoulos and Thompson 1998), offer an alternative explanation that reconciles Figure 1 with the basic mechanics in Romer. In particular, they propose that population growth expands the variety of goods (equivalently firms) and that Romer's mechanics apply within a firm, rather than in aggregate.

Peretto introduces a new variable, F, representing the endogenous number of firms. Aggregate output, Y is now defined by the number of firms, the elasticity between varieties, θ, average knowledge/productivity, A, and average employment, H_Y, per firm.

(3) $$Y = F^{\theta} A H_Y$$

Growth in knowledge conforms to Romer: $A'/A = \delta H_A$, where H_A is now average R&D labor per firm (rather than aggregate R&D labor). As in Romer, H_A equals the share of labor devoted to R&D, s, times the total labor per firm (H/F). If we assume the number of firms is proportional to employment, and define n as employment (population) growth, then steady state growth is:

(4) $$g_Y = (\theta-1)n + \delta H_A s(H/F)$$

This model eliminates the scale effects prediction, because growth from R&D is defined by the scale of the firm (captured in right hand side of equation 4) rather than by scale of the economy (captured in the left hand side of equation of equation 4). The new predictions from the model are that per capita growth increases in the share of labor in R&D, s, and in average firm size, H/F.

While this result, like Jones, solves the scale effects problem, if we assume firms reach a maximum size (defined by demand for the variety), then in the limit, R&D no longer contributes to growth—in essence a result comparable to Jones.

I propose and test an alternative interpretation of Figure 1. I argue rather than R&D getting harder, firms have become worse at it. Thus instead of challenging the scale effects assumption, I challenge the assumption that δ is fixed. This explanation preserves Romer, and accordingly the expectation of steady-state growth from R&D.

This paper empirically tests all three explanations for the disconnect between the scale effects prediction and the empirical observation in Figure 1. In testing Jones, I find that firms' maximum R&D productivity is actually increasing over time, thus providing no support for the argument that R&D has gotten harder. In testing Peretto, I find that R&D per firm is uncorrelated with per capita GDP growth—thus indicating that the Romer fails to hold even when confined to mechanics within the firm. In testing the third explanation that firms have gotten worse at R&D, I empirically find that R&D productivity has declined significantly.

To further support the argument that firms have gotten worse at R&D I present a causal (and reversible) explanation for the decline in δ. In particular I show it is due in large part to a shift in firms' R&D practices toward outsourced R&D, which has dramatically lower productivity than internal R&D (essentially zero). I also demonstrate that while the elasticity of firms' combined R&D has deteriorated over the period, there is no decline in productivity for any of the components of R&D (internal, outsourced, foreign). This provides additional evidence that R&D has not gotten harder. Finally I examine whether outsourcing is inherently less productive, or whether its lower productivity reflects selection effects: weak firms outsource and/or firms outsource weak projects. Given that the outsourcing effect persists after controlling for the selection effects, I then speculate why outsourcing has lower productivity and why, given that, firms persist with it.

This paper proceeds as follows. First I discuss the empirical approach.

Second, I discuss results. Third, I conduct deeper examination of the result that firms have become worse at R&D, by examining changes in firms' R&D practices Finally, I discuss implications.

II. Empirical Approach

This paper empirically tests each of the three explanations for the disconnect between the scale effects expectation in Romer and the empirical evidence that R&D has grown, while GDP growth has at best remained constant.

The first explanation, Jones (1995) is that R&D has gotten harder. If Jones is correct, then maximum R&D productivity should decline over time. I examine this by constructing rolling R&D productivity estimates for each firm in Compustat over 42 years (1972 to 2013). I then identify the maximum productivity in each year, and examine the significance of a time trend in that maximum. If Jones is correct, the time trend should be negative and significant.

The second explanation, Peretto (1998) is that Romer mechanics hold within the firm (the right hand side of equation 4), but that aggregate growth is muted by expansion in the number of firms (varieties) (the left hand side of equation 4). If true, then increases in R&D labor per firm should increase per capita GDP growth (because the per capita measure will ignore the left hand side). To test this, I regress per capita GDP on mean R&D labor per firm. If Peretto is correct, the coefficient on mean R&D labor should be positive and significant.

The third explanation, advanced here, is that R&D productivity has declined. To test this, I regress firm-year R&D productivity estimates on a time trend. If correct, the time trend should be negative and significant. This of course is very similar to the Jones test—in essence if R&D has gotten harder, firms will appear to have lower R&D productivity. Thus to tease apart firms getting worse from R&D getting harder, I require that maximum productivity is non-decreasing (the Jones test), while within-firm productivity is decreasing.

Thus tests for two of the three explanations require firm-specific R&D productivity estimates over time. Accordingly, our starting point in the empirics is to form those estimates. The measure we use for firms' R&D productivity is RQ (short for research quotient). RQ is the firm-specific output elasticity of R&D (Knott 2008). It is exponent γ_i in firm i's production function (Equation 5). Thus RQ matches the most common approach to measuring returns to R&D (Hall, Mairesse, Mohnen 2010). What distinguishes RQ from prior measures is that coefficients are firm-specific rather than shared across all firms in an industry.

$$Y = A_{i,}\, K_{i,t}^{\alpha}\, L_{i,t}^{\beta}\, R_{i,t-1}^{\gamma}\, S_{i,t-1}^{\delta}\, D_{i,t}^{\phi} e_{i,t} \tag{5}$$

where $Y_{i,t}$ is output, $A_{i,}$ is a firm fixed effect, $K_{i,t}$ is capital , $L_{i,t}$ is labor, $R_{i,t-1}$ is lagged R&D, $S_{i,t-1}$ is lagged spillovers, $D_{i,t}$ is advertising,

I derive RQ for each firm-year by estimating Equation 5 using a random coefficients model that allows for heterogeneity in the output elasticity for R&D (as well as all other inputs)[1]. A random coefficients model (Longford 1993) represents a general functional form model which treats coefficients as being non-fixed (across members of a cross-section or over time) and potentially correlated with the error term. Random coefficient models are those in which each coefficient has two components: 1) the direct effect of the explanatory variable and 2) the random component that proxies for the effects of omitted variables.

Use of a random coefficients (RC) specification follows from the need to capture firm specific estimates for R&D elasticity. Note that RC is a general functional form of which fixed effects is a restricted form where only the intercept is treated as a random coefficient. The random coefficients model given in Equation 6 transforms Equation 5 to ln form, and decomposes each exponent into a direct effect, $\beta_{\_}$, and the firm-specific error, $\beta_{\_i}$. Thus RQ equals $(\beta_3 + \beta_{3i})$, which corresponds to γ_i in equation 5.

$$\begin{aligned} \ln Y_{it} = {} & (\beta_0 + \beta_{0i}) + (\beta_1 + \beta_{1i}) \ln K_{it} + (\beta_2 + \beta_{2i}) \ln L_{it} + (\beta_3 + \beta_{3i}) \ln R_{i,t-1} + \\ & (\beta_4 + \beta_{4i}) \ln S_{i,t-1} + (\beta_5 + \beta_{5i}) \ln D_{it} + \varepsilon_{it} \end{aligned} \tag{6}$$

I construct RQ for each firm-year by estimating Equation 6 using rolling 7-year windows of the COMPUSTAT North American Annual database from 1965-2006. Firm level data items include (in $MM unless otherwise stated): *revenues* (Y_{it}), *capital* as net

[1] Note this discussion follows closely with that in Knott 2008.

property, plant and equipment *(K_{it})*, *labor* as full-time equivalent employees (1000)*(L_{it})*, *advertising (D_{it}),* and *R&D (R_{it})*. From these primary data, I derive a secondary measure: firm-specific *spillovers (S_{it})* which is computed as the sum of the differences in knowledge between focal firm *i* and rival firm *j* for all firms in the four digit SIC industry with more knowledge (R&D) than the focal firm:

(7) $$S_{it} = \sum_{j \neq i} R_{jt} - R_{it} \; \forall \; R_{jt} \geq R_{it}$$

This construction mimics the spillover construct in endogenous growth models with heterogeneous firms, e.g., Jovanovic and MacDonald 1994. It is a density measure that takes into account the number of firms with superior knowledge as well as the amount of each firm's surfeit of knowledge relative to the focal firm. In essence it represents the likelihood and extent of discovering superior knowledge in a random encounter with a rival firm.

I lag both R&D and spillovers one year[2], and require all firms to have a minimum of six years of data within each seven-year window. Thus each RQ estimate compares revenues to inputs using up to 7 firm-year observations. I estimate Equation 6 using Stata command, xtmixed. Xtmixed fits linear mixed models using maximum likelihood estimation. The random effects, b_i, are not directly estimated, but I form best linear unbiased predictions (BLUPs) of them using xtmixed post estimation. I then define the RQ for each firm-year as the sum $(\beta_3 + \beta_{3i})$ in the last year of each window. (See Knott and Vieregger 2014 for further details on RQ estimation, including comparisons of random coefficients and IV approaches to modeling Equation 5). A summary of Compustat data used to generate RQ is provided in Table 1.

[Insert Table 1 Here]

[2] There is no significant difference between one and two year lags. This finding matches two empirical regularities: econometric equivalence between stock and flow models and econometric equivalence of models with different lags (Griliches and Mairesse 1984, Adams and Jaffe 1996).

III. Results

3.1 Test of Jones (1995)-R&D getting harder

Results for the first test (the Jones model), are presented in Table 2. If Jones is correct that R&D has gotten harder, then maximum firm RQ should decline over time. Table 2 (Model 1) an OLS model at the firm level, reveals the opposite is true—maximum RQ is increasing over time. While only significant at the 10 percent level, the coefficient on year is positive. The coefficient estimate of 0.007 implies that max RQ is growing at roughly 7% per year. Thus, if anything, R&D appears to have become easier.

[Insert Table 2 Here]

To examine this result further, I constructed a series of industry-year maximums, where I sequentially modeled more refined definitions of industry: Model 2 examines industry defined at the 1 digit SIC code, Model 3 at the 2 digit, Model 4 at the 3 digit and Model 5 at the 4 digit. An interesting pattern emerges as industry definition becomes more refined—maximum RQ transitions from increasing over time to decreasing over time. Thus while the Jones hypothesis of R&D getting harder doesn't hold at the economy level, it may hold within industry. This result in and of itself appears to contradict the main premise in the next theory we examine—Peretto (1998). In particular, Peretto argues that endogenous growth holds within firms, but not across. Table 2 suggests the real action is across firms—more specifically across industries.

3.2 Test of Peretto (1998) - growth from R&D confined to firms

While data to test the first and third explanations come from Compustat, data for the Peretto test come from the National Science Foundation Survey of Industrial R&D (R&D labor per firm) and the World Bank (Per capita GDP growth)[3]. Because this is a simple test of one observation per firm, we present results graphically in Figure 2. If Peretto is correct that Romer holds only within firms, then increases in R&D labor per

[3] http://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG

firm should increase per capita GDP growth (because the per capita measure will consume the left hand side). Figure 2 provides little evidence of this. While correlation between the R&D labor and growth is positive, it is not significant (15.4% correlation).

[Insert Figure 2 Here]

This result that R&D per firm is uncorrelated with per capita GDP growth suggests problems with the right hand side of Equation 4. The result in Table 2 that maximum R&D productivity is increasing across, but not within industries, suggests there is also a problem with the left-hand side of equation 4. Thus there appears to be little support for the Peretto explanation.

3.3 Test of firms getting worse at R&D

To test the explanation that firms have gotten worse at R&D, I examine the trend in firm RQ. I first present this graphically in Figure 3. The figure indicates there has been a dramatic decline (65%) in mean firm RQ. Moreover, the decline coincides with the decline in implied δ from equation 1 (GDP growth divided by scientific labor).

[Insert Figure 3 Here]

While Figure 3 graphically presents evidence of declining firm productivity, I formally test that by regressing firm RQ on a time trend. Table 3 reveals that the coefficient on year is negative and significant in both OLS (Model 1) and fixed effects (Model 2). The coefficient -0.0014 in the fixed effects model indicates that on average firm R&D productivity was declining at a rate of 1.4% per year. This is consistent with the total decline in Figure 3.

[Insert Table 3 Here]

As mentioned previously, the result of declining RQ is also consistent with the Jones hypothesis—if R&D has gotten harder, firms will appear to have lower R&D

productivity. However the combination of the result here that mean firm RQ is decreasing, with that from Table 2, that maximum R&D is increasing (the Jones test), suggests that firms are getting worse at R&D, rather than that R&D is getting harder.

IV. WHAT ACCOUNTS FOR THE DECLINE IN R&D PRODUCTIVITY

The results so far provide preliminary evidence that explanation 3 (firms are getting worse at R&D) better accounts for the disconnect between Romer and the empirical evidence in Figure 1 than do the Jones and Peretto explanations. However, the result that firms have gotten worse would be more compelling if we could tie it to firm behavior. Doing so requires data on firms' R&D practices. Thus while the data for the primary tests come from publicly available sources, data for deeper examination comes from the National Science Foundation (NSF) Survey of Industrial Research and Development (SIRD). The SIRD is an annual survey of firms conducting R&D in the US. Since historically 69% to 75% of US R&D is performed by firms[4], this comprises the majority of US R&D activity. The SIRD was collected through a joint partnership between the sponsoring agency, the National Science Foundation (NSF), and the Census Bureau from 1957 to 2007 (though microdata is only available from 1972 forward)[5].

The SIRD is gathered from a sample intended to represent all for-profit R&D-performing companies, either publicly or privately held. (see Foster and Grimm 2010 for details). The sampling methodology comprises a census of all firms whose R&D expenditure exceeds one million dollars, and a random sample of firms with smaller budgets. In order to minimize the burden of completing the survey, the SIRD collects some data each year, and other data in alternating years. The data collected annually includes: domestic sales, domestic employment, number of scientists/engineers, and total domestic R&D expenditures by source of funding (Federal R&D funds versus company R&D funds), horizon (basic research, applied research, and development), and location (internal, outsourced within US, conducted by foreign entities). The data collected in

[4] www.nsf.gov/statistics/seind12/c4/c4s1.htm
[5] As of 2008, the SIRD was substantially redesigned and renamed the Business Research and Development and Innovation Survey (BRDIS).

alternating years adds: total R&D costs decomposed by type (wages and salaries of R&D personnel, costs of materials, depreciation on R&D property and equipment, and other costs). For R&D conducted internally, it also includes distribution across the 50 states and the District of Columbia.

Survey instructions provide details on what to include in each category (as well as give examples). For example, R&D includes: *Basic research*: "the planned, systematic pursuit of new knowledge or understanding toward general application", *Applied research:* "the acquisition of knowledge or understanding to meet a specific recognized need", and *Development*: "the application of knowledge or understanding toward the production or improvement of a product, service, process or method". However, R&D explicitly excludes "quality control, routine product testing, market research, sales promotion, sales service, routine technical services and other nontechnological activities."

I construct a variable *internal R&D,* by summing three SIRD variables (basic research, *brtot*, applied research, *ardtot*, development, *devtot*). Other variables include: US outsourced R&D, *outuscomp*, and foreign R&D, *outforeign)*, domestic revenues (*dns*), the number of full-time equivalent *employees (dne)*. A summary of these variables is presented in Table 4.

[Insert Table 4 Here]

The first step in understanding the decline in RQ examines trends in the SIRD data. Those data reveal that the only trend as dramatic as the decline in RQ is the rise in R&D outsourcing (Figure 4). Mean firm outsourcing increased by a factor of 20.5 for the thirty-year period 1972 to 2001. In contrast, foreign R&D grew by a factor of 3.0 and scientific labor by a factor of 2.4. This prompts the hypothesis that outsourcing may have contributed to the decline in δ.

[Insert Figure 4 Here]

This shift from internal R&D to outsourced R&D is only relevant to total decline in δ however, if the productivity of internal and outsourced R&D differ. To test that, I decompose R&D investment into its three SIRD constituents (internal, outsourced, and foreign), then combine these constituents with other inputs in the firm's production function (Equation 5). I then estimate the contribution of each form of R&D to firm output. Note that SIRD does not collect data on capital or advertising, thus this estimation suffers from omitted variables. Comparison of RQ estimates of Compustat data with the full set of variables in equation 5 versus the abbreviated set of SIRD variables, indicates the RQs from the two approaches are correlated at 92%.

Results are presented in Table 5. Model 1 present results for the full period (1983 to 2000).[6] The model indicates the elasticity of internal R&D is 0.13, whereas the elasticity of outsourced R&D is essentially zero (0.001). This means a 10% increase in internal R&D increases revenues 1.3%, whereas a 10% increase in outsourced R&D has no impact on revenues. Since the scales of internal versus outsourced R&D differ, we also evaluate the marginal values in dollars. An additional $100,000 on internal R&D at the mean yields a 1.3% increase in revenues. In contrast an additional $100,000 in outsourced R&D yields a 0.03% increase in revenues—a 43X difference. These results are robust to the exclusion of spillovers (Model 2). Thus outsourcing appears to be unproductive for the funding firm.

[Insert Table 5 Here]

This result seems fairly compelling, but because I posit that outsourcing rather than "R&D getting harder" is responsible for the decline in RQ, it is useful to examine how coefficients on the R&D elements are changing over time. To do so, I subdivide the data from Model 1 into three periods, and re-estimate the production function for each period (Models 3-5). Results indicate the elasticities for all three components of R&D are fairly stable over time. Thus, there is no support for the hypothesis that R&D has gotten harder. More interestingly, firms have not gotten worse at any element of R&D. Thus the decline in RQ appears to stem from reallocation of R&D resources from more

[6] Prior to 1983 not all SIRD variables were available all years; post 2000 some variable definitions changed

productive (internal) to less productive (outsourced) elements.

The result that outsourcing is unproductive raises the question of whether the effect is real or merely an artifact of who outsources (firm quality) or what gets outsourced (project quality). Tackling the firm quality question requires some understanding of why firms outsource. Interviews with R&D managers indicate they outsource for a number of reasons, and that those reasons vary with top-level R&D strategy rather than industry mandate (in less than 5% of industries is it the case that all firms outsource). At one end of the outsourcing spectrum, firms outsource only to universities and government labs. They do this to gain access to basic research as well as to identify potential employees. In the middle of the spectrum, firms outsource under special circumstances. For example, they use outsourcing as a flexible substitute for internal hiring when future demand is uncertain; they outsource activities where they lack capability and don't intend to develop it internally because they would operate below efficient scale; or they outsource testing (particularly in the case of pharmaceutical trials by contract research organizations, CROs). At the extreme end, firms outsource all non-core R&D activities.

The SIRD did not collect data on the destination for outsourced R&D, however these data were collected in its successor, the Business R&D and Innovation Survey (BRDIS). A report of BRDIS data (Morris and Schakelford 2014) indicates 3.4% of outsourcing is to universities, 81.3% is to companies, and the remaining 15.2% is to government agencies and other organizations. Thus the vast majority of outsourcing is to firms, suggesting it conforms to rationale in the middle and extreme ends of the spectrum.

To test whether firm quality is driving the outsourcing result, I construct a two-stage treatment model of the impact of outsourcing, where the first stage models the firm's decision to outsource, and the second stage models the treatment effect of that choice on internal R&D productivity. To estimate the first stage, I construct a variable, *neverout*, which captures the firm's outsourcing decision. *Neverout* takes on the value 0 if the firm has ever outsourced, and takes on the value 1 otherwise. Because of the limited number of variables in SIRD, no first stage model was able to explain more than 5% of the variance in *neverout*. Despite that, the treatment model (Table 6) yields interesting results.

[Insert Table 6 Here]

Model 1 captures the impact of outsourcing on *internal RQ*. If quality of firms who outsource drives the lower productivity of outsourcing, the coefficient for *neverout* should be significant in the second stage model of internal R&D productivity. This "treatment effect" would mean that firms who outsource have lower internal R&D productivity. Table 6 indicates that's not the case. The coefficient on *neverout* in model 1 is zero. This implies that outsourcing firms are of the same quality as non-outsourcing firms (have the same internal R&D productivity). In contrast, Model 2 repeats the test but examines the treatment effect of outsourcing on the *total RQ* (the sum of internal, outsourced and foreign). Here the coefficient on *neverout* is significant. Because the first test rules out differences in internal RQ, the second test suggests the lower RQ of outsourcing firms stems from outsourcing itself.

While firm quality doesn't appear to drive the outsourcing result, it is still possible the lower productivity of outsourced R&D is driven by project quality—firms outsource their lower quality projects. To investigate that, I examined what happens to firms' internal R&D productivity surrounding the time they first outsource. I do this by creating an eleven-year event window around the time a firm first outsources R&D (year zero $\pm$5). I then regress moving estimates of firms' internal RQs on the set of relative-year dummies. If firms are merely outsourcing their lower quality R&D projects, their internal productivity should increase immediately after they shunt those projects to outsourcing. Table 7 indicates that doesn't happen. In fact, the mean coefficient prior to outsourcing (0.0006) is actually higher than post-outsourcing (0.0000), though not significantly. Thus it appears unlikely project quality is driving the lower productivity of outsourcing.

[Insert Table 7 Here]

Since neither low quality firms nor low quality projects explains the lower productivity of outsourced R&D, it appears outsourced R&D inherently has lower productivity than internal R&D. What might drive that? While the data don't provide insight, a reasonable explanation is that R&D produces internal spillovers. We know this to be true across firms, e.g., Spence 1984, so it seems reasonable this would be true within firms. To understand the intuition for internal spillovers, consider the fact that firms carry portfolios of projects. The most likely outcome from each project is termination: a survey of member firms in the Industrial Research Institute (IRI) indicates reports that on average only two in 125 funded projects are ultimately launched commercially (Stevens and Burley 1997). This means the value of R&D projects lies outside project outcomes themselves—most likely in the ability to recycle knowledge gleaned from the failures into subsequent projects.

A specific example of this from Hughes Aircraft Company (Chester 1994) pertains to ion propulsion technology. The technology was originally developed for attitude correction of military satellites, though its real advantage lay in long-term reliability. Given military satellites generally have missions lasting only five years, the higher upfront costs of ion propulsion were never warranted, so military development was terminated. However because of Hughes' broad product base, it was able to redeploy the technology for previously unforeseen applications such as ion implantation for semiconductor fabrication. While the Hughes case is anecdotal, a recent study of banks' adoption of internet banking finds that firms who outsource the initial IT integration are less able to develop new applications and accordingly have lower revenues from their internet operations (Weigelt 2009).

An alternative explanation for inherently lower productivity of outsourced R&D is that it is more costly to exploit and/or less likely to be exploited because key technical resources lie outside the firm. This explanation would be consistent with early work by Tom Allen (1977) showing that R&D project efficiency is driven in part by proximity, and with later work (Clark and Fujimoto 1991) showing that co-locating design engineers and manufacturing engineers dramatically reduced the duration and cost of automobile development. Similar effects have been found for technology outsourcing (Helfat and Raubitschek 2000, Weigelt 2009).

In summarizing this deeper examination of the decline in firms' R&D productivity, it appears to be driven by a shift toward greater used of outsourced R&D. this conclusion is driven by the facts that: a) Outsourcing has grown at ten times the rate of R&D labor, and b) outsourced R&D is unproductive for the funding firm. Moreover, additional tests indicate the lower productivity of outsourcing appears to be fundamental, rather than an artifact of firm quality or project quality.

V. Discussion

Since at least Solow (1957) there is widespread belief that technological progress drives economic growth. The first theory to explain the mechanism through which R&D investment drives growth (Romer 1990) yielded the "scale effects" prediction that growth should increase in the level of scientific labor. However empirical evidence conflicts with that: scientific labor has been increasing, while GDP growth has been declining.

Two theories have been advanced to explain the disconnect between theory and empirical evidence. The first (Jones 1995) holds that R&D has gotten harder: the second (Peretto 1998) holds that Romer applies within the firm, rather than in aggregate. Since both theories imply zero growth from R&D in steady-state (which is inconsistent with the spirit of endogenous growth theory), I offered a third explanation that preserves the Romer model: R&D productivity has declined.

I tested all three explanations. With regard to Jones, I found no evidence that R&D has gotten harder. Rather I found that maximum RQ has increased, while internal RQ has remained constant. With regard to Peretto, I found there was no significant correlation between mean R&D labor and per capita growth. Thus there was no evidence that Romer scale effects hold even within the firm. With regard to the third explanation, I found that firms' R&D productivity (RQ) had declined dramatically. Moreover the decline in firm RQ was highly correlated with the decline in Romer's δ. While this could be due to R&D getting harder (Jones) or firms getting worse at it, the further finding that internal RQ had remained constant ruled out Jones. Taken together, these results seemed to offer considerable support for the third explanation, but no support for the alternative explanations.

The bulk of the analysis then explored the "getting worse" explanation further. I found first, that the composition of firms' R&D had changed. In particular, outsourced R&D had increased 20.5x during a period over which R&D labor had increased 2.5x.

I found second, that outsourced R&D was unproductive--its output elasticity was essentially zero. Thus it appears outsourced R&D offers a plausible explanation for the decline in R&D productivity. Further, tests of firm quality and project quality suggest that neither is driving the lower productivity of outsourced R&D. Thus outsourced R&D appears to be inherently less productive.

While the question of why outsourced R&D has lower productivity than internal R&D is interesting and important, a related and equally important question is why firms persist with outsourcing, given its lower productivity. Again the SIRD data provide no insights. However there is substantial evidence that firms don't know the underlying productivity of their R&D (Knott 2009). Indeed the Industrial Research Institute (IRI) reports the need for better R&D metrics is a top concern of members (Schwartz, Miller, Plummer, Fusfeld 2011).

Given this ambiguity about the productivity of R&D, firms are vulnerable to information cascades. An information cascade occurs when it is optimal for actors to follow the behavior of preceding actors without regard for their own information (Bikchandani, Hirschleifer and Welch 1992). The relevant cascade in this instance pertains to open innovation, which was first mentioned in 1983, but was later popularized by books (Chesbrough 2000) as well as articles in MIT Sloan's Review (Chesbrough 2003) and Business Week (Eingardio and Einhorn 2005). The espoused benefits of open innovation include reduced cost of R&D as well higher R&D productivity. Evidence of the power of the cascade comes from a survey of CIOs and CEOs which reveals 70% of them believe outsourced innovation improves financial performance (Oshri and Kotlarsky 2011), which is unlikely given the evidence presented here.

It is worth noting that outsourcing is just one form of open innovation. Another form—"external knowledge sourcing" appears to increase innovation (Cassiman and Veugulers 2006, Arora, Cohen and Walsh 2014). The key distinction between the two forms of open innovation is that external knowledge sourcing pertains to the locus for obtaining an idea, whereas R&D outsourcing pertains to the locus for researching and developing an idea.

While the questions of why outsourcing has lower productivity and why, despite that, firms persist with it, are ripe areas for future research, we have made progress in explaining the disconnect between Romer's theory and the empirical record. It is not that R&D has gotten harder or that Romer's mechanics are confined to the firm, both of which yield zero growth in steady-state. Rather it appears firms have become worse at it--their R&D productivity has declined. This decline stems from reallocation of R&D resources from more productive (internal R&D) to less productive (outsourced R&D) activity, rather than from fundamental decay in R&D capability.

These results have two implications. First, Romer's prospect of steady-state growth from R&D may still be valid. Second, it may be possible to restore R&D productivity and growth fairly simply—by gradually bringing outsourced R&D back inside the funding firm.

REFERENCES

Aghion, P and P. Howitt 1992. “A Model of Growth through Creative Destruction”, *Econometrica* 60: 323-351.

Aghion, P and P. Howitt 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press

Allen, T. 1977. *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information Within the R&D Organization*. Cambridge, MA: MIT Press

Arora, A., W. Cohen, and J. Walsh (2014) “The Acquisition and Commercialization of Invention in American Manufacturing: Incidence and Impact”, *NBER Working Paper* No. 20264

Beck, N. and J. Katz. 2007. Random coefficient Models for Time-Series-Cross-Section Data: Monte Carlo Experiments, *Political Analysis* 15: 182-195.

Bikhchandani, S., D. Hirshleifer, D., and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," Journal of Political Economy, 100(5): 992-1026.

Cassiman, B and Veugulers (2006) “In search of complementarity in innovation strategy: internal R&D and external knowledge acquisition” Management science 52 (1), 68-82.

Chesbrough, H. (2003). "The era of open innovation". MIT Sloan Management Review 44 (3): 35–41

Chesbrough, Henry William (1 March 2003). Open Innovation: The new imperative for creating and profiting from technology. Boston: Harvard Business School Press

Chester,A. 1994. "Aligning Technology with Business Strategy", *Research and Technology Management* , Jan-Feb 1994:25-32.

Clark, K and T. Fujimoto (1991) Product Development Performance: Strategy, Organization, and Management in the World Auto Industry Harvard Business Review Press.

Dinopoulos, E. and P. Thompson 1998. “A Schumpeterian model of protection and relative wages” .*Journal of Economic Growth* 3 (4), 313-335.

Eingardio, P and B. Einhorn (2005) “Outsourcing Innovation”, Business Week, March 20.

Foster, L. and C. Grimm (2010) "Characteristics of the Top R&D Performing Firms in the U.S.: Evidence from the Survey of Industrial R&D." *PDF Center for Economic Studies Discussion Paper Series*, No. 10-33.

Grossman, G. and E. Helpman 1991. Quality Ladders in the Theory of Growth” *Review of Economic Studies*, 58: 43-61.

Hall, B. and J. Mairesse and P. Mohnen, 2010. “Measuring the Returns to R&D”, Handbook of the Economics of Innovation, B. H. Hall and N. Rosenberg (eds.), North-Holland.

Helfat, C. and R. Raubitscheck 2000. “Product sequencing: co-evolution of knowledge, capabilities and products” Strategic Management Journal, 21(10-11): 961–979

Jones, C. 1995. “R&D-Based Models of Economic Growth”, Journal of Political Economy, 103(4): 759-784

Jovanovic, B. and G. MacDonald 1994. "The Life Cycle of a Comperitive 1ndustty”, *Journal of Political Economy* 102(2): 322—47

Knott, A.M. 2008. “R&D/Returns causality: Absorptive capacity or organizational IQ”. Management Science, 54, 2054–2067

Knott, A.M. and C. Vieregger 2014 "An Empirical Test of Endogenous Firm Growth" *SSRN Working paper* 2382885.

Knott, Anne Marie (2009) "New Hope for Measuring R&D Effectiveness", *Research-Technology Management* 53(5) 9-13

Kortum, S. 1997 “Research, patenting, and technological change”.*Econometrica* 65 (6), 1389-1419.

Longford, N. 1993. *Random Coefficients Models*, New York: Oxford University Press

Moris, F and B. Shackelford 2014, “Extramural R&D Funding by U.S.-Located Businesses Nears $30 Billion in 2011”. *National Science Foundation, National Center for Science and Engineering Statistic*s, NSF 14-314.

Oshri, I and J. Kotlarsky (2011). Innovation in Outsourcing: A Study of Cleint Expectations and Commitment. Warwick Business School working paper. *Research-Technology Management*, 40(3):

Peretto, P. 1998. “Technological Change and Population Growth“, *Journal of Economic Growth*, 3(4): 283-311.

Romer, P. 1990. “Endogenous Technological Change” *Journal of Political Economy* 98(5):S71-S102.

Schwartz, L., R. Miller, D. Plummer, and A. Fusfeld, (2011) “Measuring the Effectiveness of R&D”, Research-Technology Management, 54(5): 29-36.

Segerstrom, P. 1998. “Endogenous Growth without Scale Effects”, *American Economic Review*, 88(5): 1290-1310.

Solow, Robert M (1957). "Technical Change and the Aggregate Production Function". Review of Economics and Statistics (The MIT Press) 39 (3): 312–320

Stevens, G. and J. Burley.1997. “3,000 Raw Ideas Equals 1 Commercial Success!” *Research-Technology Management*, 40,(3) :

Weigelt , C. (2009). The impact of outsourcing new technologies on integrative capabilities and performance, Strategic Management Journal 30, 595-616.

Young, A. “Growth without Scale Effects.” *Journal of Political Economy* 106: 41-63

TABLE 1. Compustat Data Summary

Observations=55947

	Mean	StDev	1	2	3	4	5	6	7
1. l(revenue)	5.24	2.47	1.00						
2. l(capital)	3.42	2.92	0.95	1.00					
3. l(employees)	0.19	2.30	0.95	0.93	1.00				
4. l(R&D)	1.89	2.43	0.82	0.75	0.74	1.00			
5. l(advertising)	-5.07	5.31	0.19	0.16	0.20	0.18	1.00		
6. l(spillovers)	5.80	4.18	-0.17	-0.21	-0.28	0.04	-0.05	1.00	
7. RQ	0.10	0.08	0.16	0.13	0.15	0.18	0.07	-0.06	1.00

TABLE 2. Test of Jones hypothesis that R&D has gotten harder

	1 OLS	2 FE	3 FE	4 FE	5 FE
Dependent variable	yearRQmax	sic1yr RQmax	sic2yr RQmax	sic3yr RQmax	sic4yr RQmax
Year	0.007^	**0.002**	0.000	**-0.001**	**-0.001**
	0.004	0.001	0.000	0.000	0.000
Industry effects		included	included	included	included
Constant	-13.900	**-4.245**	0.126	**1.961**	**2.138**
	8.731	1.514	0.449	0.206	0.161
R-squared	0.067	0.011	0.000	0.010	0.012
within		0.023	0.000	0.013	0.016
between		0.289	0.012	0.000	0.001
adjusted R-sq	0.043				
observations	41	402	1938	6197	9627
groups (sic)		10	65	235	367

TABLE 3. Test of Firms getting worse at R&D

	1	2
	OLS	FE
Dependent variable	RQ	RQ
Year	**-0.0009**	**-0.0014**
	0.0000	0.0000
firm effects		yes
Constant	**1.9233**	**2.8637**
	0.0660	0.0853
R-squared	0.0134	0.0134
within		0.0206
between		0.0019
adjusted R-sq	0.0134	
observations	55957	55957
groups (firms)		6217

TABLE 4. NSF Survey of Industrial R&D (SIRD) Data

28500 Firm-year Observations (3500 firms)

	Mean	Std.Dev.	1	2	3	4	5	6
1.ln(revenue)	10.89	2.28	1.00					
2.ln(scientists)	2.82	1.93	0.53	1.00				
3.ln(employees)	6.20	2.00	0.92	0.76	1.00			
4.ln(InternalRD)	7.73	2.46	0.63	0.91	0.71	1.00		
5.ln(outUS)	5.45	2.47	0.29	0.54	0.33	0.63	1.00	
6.ln(foreign)	7.27	2.29	0.44	0.62	0.49	0.64	0.51	1.00

TABLE 5. Decomposing R&D productivity

Random coefficients estimation	(1)	(2)	(3)	(4)	(5)
Dependent variable: ln(Revenues) *Yit*	1983-2000	1983-2000	1983-1988	1989-1994	1995-2000
ln(employees) *(Lit)*	**0.841**	**0.843**	**0.885**	**0.867**	**0.921**
	0.006	0.006	0.018	0.015	0.023
ln(internalR&D) *(Rit-1)*	**0.128**	**0.126**	**0.114**	**0.094**	**0.108**
	0.004	0.004	0.015	0.010	0.011
ln(outsource R&D) *(Ot-1)*	0.001	0.002	**0.008**	**0.006**	**0.009**
	0.002	0.002	0.003	0.003	0.004
ln(foreignR&D) *(Fit-1)*	**0.008**	**0.007**	0.003	0.004	0.000
	0.002	0.002	0.004	0.004	0.003
ln(spillovers) *(Sit-1)*	**0.004**		0.003	-0.005	-0.008
	0.001		0.002	0.002	0.002
Constant	**4.872**	**4.915**	**4.487**	**5.210**	**4.931**
	0.049	0.047	0.147	0.113	0.123
Log-likelihood	-25911	-26195	-2094	-5019	-3902
Wald chi2	23438	24022	3939	5649	8292
prob>chi2	0	0	0	0	0
Observations	28500	28500	2500	5000	5000
Firms	3500	3500	500	1000	1000

Obs and firms rounded to nearest 500
Std errors below coefficients
Coefficients in bold significant at 0.05

TABLE 6. Treatment test of firm quality

Treatment regression	(1)	(2)
Dependent variable:	Internal RQ	Aggregate RQ
ln(employees)	0.000	**-0.004**
	0.000	0.000
Scientist ratio (dns/dne)	**0.006**	**0.012**
	0.001	0.002
Scientist cost (totcost/dns)	**0.000**	**0.000**
	0.000	0.000
Neverout	0.000	0.015
	0.001	0.008
Constant	**0.093**	**0.087**
	0.001	0.007
wald chi2	40.46	484.5
prob>chi2	0	0
stage 1 neverout		
Basic percent (brtot/rdtotown)	**0.511**	**0.508**
	0.122	0.117
Applied percent (ardtot/rdtotown)	-0.186	**-0.291**
	0.124	0.118
Foreign percent (outforeign/RDtrue)	**-1.330**	**-1.451**
	0.215	0.194
Federal percent (fedtot/Rdtrue)	0.194	0.098
	0.164	0.180
Constant	**0.857**	**0.870**
	0.078	0.073
/ath rho	0.030	-0.242
/lnsigma	-4.548	-3.676
rho	0.030	-0.237
sigma	0.010	0.025
lambda	0.000	-0.006
prob>chi2	0.690	0.026
observations	6000	6500
firms	500	500

Obs and firms rounded to nearest 500
Std errors below coefficients
Coefficients in bold significant at 0.05

TABLE 7. Test of Project Quality

	(1)
Dependent variable: Internal RQ	
t-5	**0.003**
	0.001
t-4	0.002
	0.001
t-3	**0.003**
	0.001
t-2	**-0.004**
	0.001
t-1	-0.001
	0.001
first year of outsourcing	**-0.003**
	0.001
t+1	**0.002**
	0.001
t+2	**-0.002**
	0.001
t+3	**0.004**
	0.001
t+4	**-0.003**
	0.001
t+5	**0.001**
	0.000
Constant	**0.097**
	0.000
R-squared	0.013
Adjusted R-squared	0.012
Observations	15000
Firms	1000

Obs and firms rounded to nearest 500
Std errors below coefficients
Coefficients in bold significant at 0.05

FIGURE 1. Empirical Evidence Questioning Romer's Scale Effects

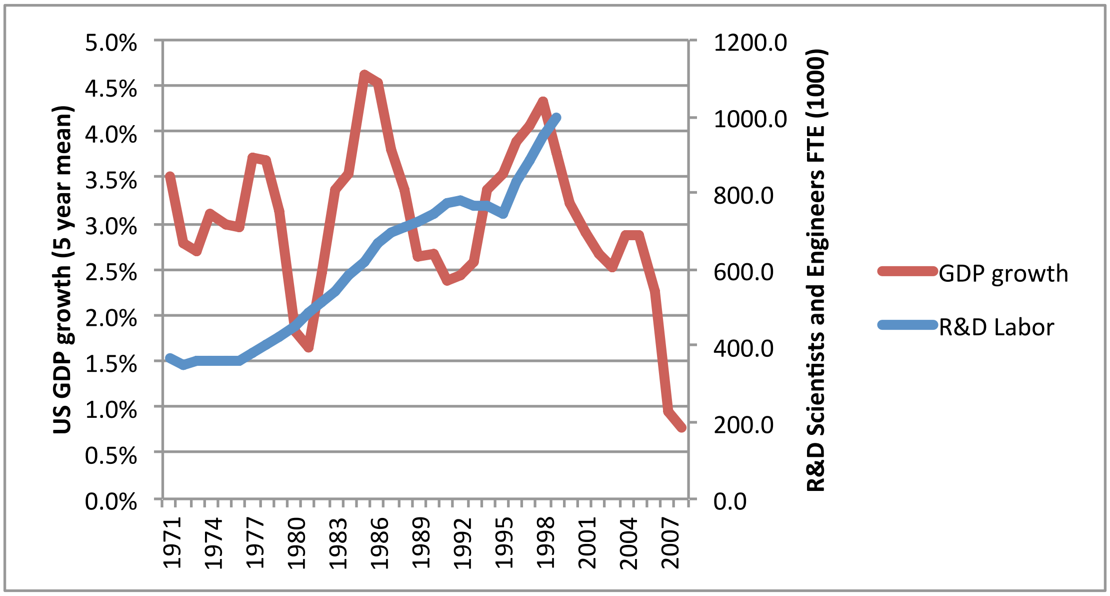

FIGURE 2 Test of Peretto

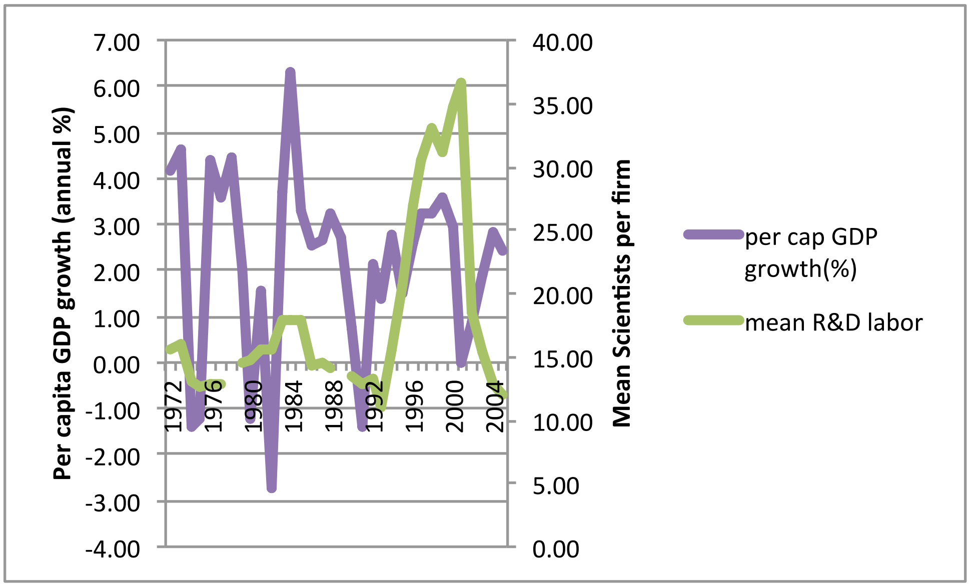

FIGURE 3. Test of R&D getting harder

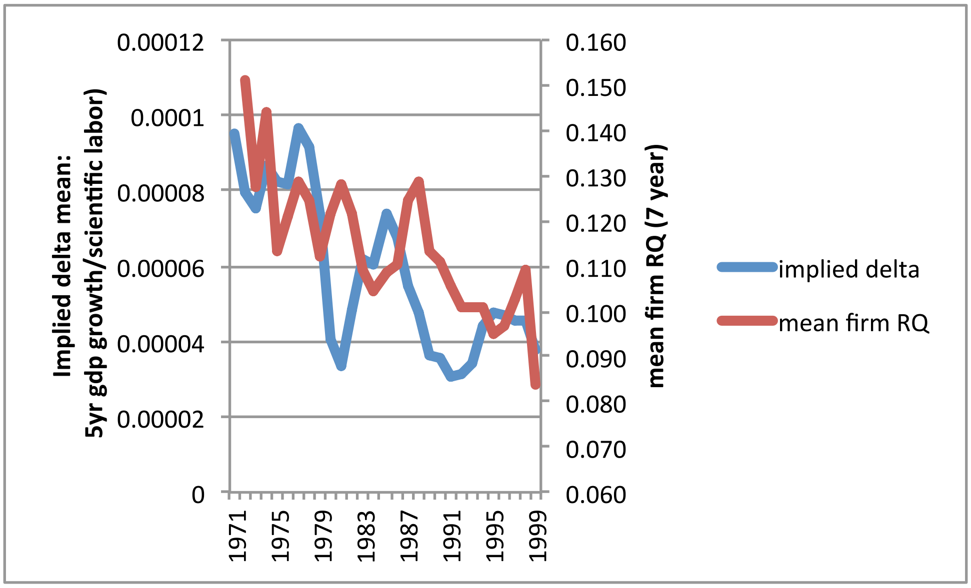

FIGURE 4. Trends in SIRD Data