"Why Are Some Firms More Innovative? Knowledge Inputs, Knowledge Stocks, and the Role of Global Engagement"*

Chiara Criscuolo
University College London

Jonathan E. Haskel
Queen Mary, University of London,
Fellow, UK Advanced Institute of Management,
and CEPR

Matthew J. Slaughter
Tuck School of Business at Dartmouth and NBER

Initial Draft: March 2004

JEL Classification: F1, F2, 03
Key Words: Multinational Firms, Exporting, Knowledge and Technological Change

Abstract

Why do some firms create more knowledge than others? This question is typically answered in macro and industrial-organization literatures with reference to a production-function model in which new ideas spring from the interaction of researchers and the existing stock of knowledge. But there is very little empirical evidence on production functions for new ideas. In this paper we estimate knowledge production functions for a cross-section of U.K. firms covering their operations from 1998 through 2000. We focus in particular on the hypothesis from the trade literature that globally engaged firms—either multinationals or exporters—have access to larger knowledge stocks. We find that globally engaged firms do generate more ideas than their purely domestic counterparts. This is not just because they use more researchers. Importantly, it is also because they have access to a larger stock of ideas through two main sources: their upstream and downstream contacts with suppliers and customers, and, for multinationals, their intra-firm worldwide pool of information.

---

*Email addresses: Chiara.Criscuolo@ons.gsi.gov.uk; j.e.haskel@qmul.ac.uk or Jonathan.Haskel@ons.gsi.gov.uk; matthew.j.slaughter@dartmouth.edu.

For financial support Haskel thanks the U.K. Economic and Social Research Council, and Slaughter thanks the National Science Foundation.

1. Introduction

A crucial determinant of economic growth is the creation and spread of new knowledge. In the macro literature on how countries grow, a central role is played by technological advance via the flow of new ideas (e.g., the survey of Jones, 2003). This central role also appears at the micro-level in the industrial-organization (IO) literature on how firms become more productive (see e.g., Griliches, 1998).

Much of the theory in these literatures models new ideas as coming from two main sources. One is investment in researchers and scientists. The other is the stock of existing knowledge, which is, at least to some extent, a public good. Thus, new knowledge depends on people working at discovery and the stock of existing knowledge, a knowledge or ideas production-function model that is usually credited to Griliches (1979).

In the macro growth literature, this public-good feature is often interpreted to mean there exists a single global stock of knowledge equally accessible to all actors around the world. For example, Jones (2002) models a framework in which "Ideas created anywhere in the world are immediately available to be used in any economy. Therefore, the [stock of ideas] used to produce output … corresponds to the cumulative stock of ideas created anywhere in the world and is common to all economies." Many other studies use similar frameworks, e.g., Parente and Prescott (1994), in which "World knowledge … is meant to represent the stock of general and scientific knowledge in the world (i.e., blueprints, ideas, scientific principles, and so on). We assume that all firms have access to this knowledge. Thus, general and scientific knowledge spills over to the entire world equally."

Although there is widespread acceptance of this framework of a production function for new ideas, there is very little evidence on its empirical validity. Indeed, in his forthcoming literature survey Jones (2003) puts this problem at the top of his concluding "to do" list for future research.

> While we have made much progress in understanding economic growth in a world where ideas are important, there remain many open, interesting research questions. The first is, 'What is the shape of the ideas production function?' How do ideas get produced? … The current research practice of modeling the ideas production function as a stable Cobb-Douglas combination of research and the existing stock of ideas is elegant, but at this point we have very little reason to believe that it is correct.

Knowing how much economic entities (firms, industries, countries) benefit from knowledge stocks outside their own boundaries is difficult in part because of measurement: measuring internally chosen inputs for creating knowledge should be relatively easy (e.g., number of scientists), but by construction measuring both stocks of external knowledge and the extent to which they flow across entities is harder.

At the country level, one reason to suppose that more empirical work is needed is that so much of the large cross-country variation in per capita income is accounted for by variation in the level of the Solow residual, rather than in the endowments of capital (Hall and Jones, 1999). It seems conceivable that large cross-country differences in measured Solow residuals might at least partly reflect cross-country differences in the creation of and/or access to ideas.

What in the real world might help explain differences in creating and/or accessing ideas? The international trade literature suggests an answer: the global engagement of firms. In recent years a growing body of empirical evidence indicates that firms that are multinational and/or exporters are particularly "knowledge intensive."

For example, in the manufacturing sectors of the United States (Doms and Jensen, 1998) and the United Kingdom (Criscuolo and Martin, 2003) it is multinational firms—parents and also affiliates of foreign-owned firms—that show the highest levels of total factor productivity (TFP). Similarly, exporters in many countries exhibit high productivity levels and/or growth (see the survey of Tybout, 2000). There is also evidence that multinationals use more knowledge inputs: e.g., multinationals seem to do more advertising (Brainard, 1997), or the evidence in Bernard, Knetter, and Slaughter (2004) that in recent decades the parents of U.S.-based multinationals have consistently performed about two-thirds of all U.S. private-sector research and development (R&D) despite accounting for barely $1/20^{th}$ of 1% of all firms.

Motivated by this body of empirical evidence, the now-standard trade models of multinationals (Markusen, 2002 for an extensive treatment, or Carr, et al, 2001 for an abridged summary) make the crucial assumption that these firms are particularly knowledge intensive relative to purely domestic firms. Indeed, in this "knowledge capital" model multinationals arise via foreign direct investment (FDI) largely because of the desire (and ability) to deploy firm-specific knowledge assets in multiple countries despite the co-ordination and set-up costs of multi-plant production. There is ample evidence of this cross-border intra-firm knowledge transfer (e.g., Mansfield and Romeo, 1980, and Moran, 2001). There is mixed evidence, at best, whether this knowledge somehow "spills over" from affiliates to domestic firms in host countries (e.g., Aitken and Harrison, 1999, and Haskel, et al, 2002).

One potential limitation of this empirical evidence from the international literature is that it does not map neatly into the growth/IO framework of the ideas production function. Yes, multinationals seem to use more knowledge inputs like R&D scientists than do purely domestic firms. Does that fact alone account for the evidence from various countries of their higher TFP? We know of no systematic examination of the alternative story that access to knowledge stocks might vary across the global engagement of firms. But additional facts suggest that examining these alternative stories from the growth/IO perspective would be fruitful: e.g., the fact that the handful of highest-income countries noted by macroeconomists (e.g., in Hall and Jones, 1999) are also the countries that source and receive the large majority of world trade and FDI (e.g., Markusen, 2002) and R&D (e.g., Keller, 2001).

The goal of this paper is to apply the ideas production function from the macro and IO literatures to a data set of globally engaged firms suggested by the trade literature. Specifically, we estimate knowledge production functions for a cross-section of U.K. firms covering their operations from 1998 through 2000. These data come from the EU-wide Community Innovations Survey (CIS); for each firm sampled we have detailed measures of knowledge outputs, knowledge inputs, and sources of knowledge. These measures of sources of knowledge are a particular innovation in our study: in our data firms report sources of learning from a wide range of market and non-market options. We also know whether each firm is globally engaged either in terms of being either an exporter (information recorded directly in the CIS) or part of a multinational firm (information we merged into the CIS).

Our estimates of various specifications of knowledge production functions yield three main empirical results. First, globally engaged firms do in fact generate more knowledge output than do domestic firms. This is true for several alternative measures, including patents filed or whether the firm undertakes process or product innovation. Second, the first fact is explained partly by the fact that globally engaged firms employ more knowledge inputs. Again, this is true for several alternative measures such as R&D scientists or total expenditures on knowledge.

Third and most importantly, our estimates contradict the hypothesis that all firms have access to the same stock of worldwide knowledge. Instead, globally engaged firms appear to generate more knowledge output thanks to access to a larger stock of ideas through two main sources:

their upstream and downstream contacts with suppliers and customers; and, for multinational firms (both parents of U.K.-headquartered firms and affiliates of foreign-owned firms), their intra-firm worldwide pool of information.

Taking our three findings together, we conclude that globally engaged firms are more innovative both because they use more knowledge inputs and also because they have access to larger knowledge stocks. For the growth and IO literatures whose techniques we are applying, the most important message might be there is no single global stock of knowledge to which all firms have equal access. An important determinant of the aggregate stock of knowledge to which a country has access may thus be how globally engaged are its firms via trade and FDI. This possibility has been acknowledged in these literatures (e.g., the comment in Jones, 2003, that "Intuitively, openness to international trade is likely related to openness to idea flows"). Our findings offer clear evidence on these sorts of conjecture.

There are four sections to the rest of the paper. Section 2 briefly discusses the theory of producing knowledge. Section 3 discusses our data, measurement, and estimation issues. Section 4 presents our empirical findings, and Section 5 briefly concludes.

2. A Theoretical Framework of Knowledge Production

A Basic Production Function

Like an output production function, the innovation production function relates inputs into the innovation process to outputs. The classic form of the innovation production function in works such as Griliches (1979) and Romer (1990) can be written as follows.

$$\Delta A_i = AH_i \qquad (1)$$

where ΔA is the change in knowledge stock—i.e., the flow of new ideas; A is the initial stock of knowledge; and H is inputs into the process of knowledge creation—e.g., the human capital of R&D scientists. I is used to index variously countries, industries, or firms; for our study, it will index firms within a country (the United Kingdom).

As Romer (1990) stresses, equation (1) makes important assumptions about mechanisms and functional form. First, increases in knowledge depend on the knowledge stock researchers have to work with: the "standing on shoulders" effect. Thus, (1) says that 100 researchers working today produce more knowledge than 100 researchers working a century ago, thanks to the fact that today's researchers have more A to work with.

Second, not subscripting A on the right hand side of (1) means that scientists in unit I all have access to the same knowledge stock. As discussed in the introduction, many researchers have modelled this to mean a single worldwide stock of knowledge to which all firms have equal access. Testing this assumption empirically will be a central feature of our empirical work.

Third, equation (1) has a multiplicative functional form with no diminishing returns to A or H. This assumption has been much debated in the literature; see Jones (2003) for an overview. Because many of the preferred innovation measures in our data are categorical, we will not be able to estimate parameters of particular functional forms.

Ignoring issues of functional form, consider the implicit assumption that all knowledge is a public good. As discussed in the introduction, this assumption might not accord with the trade idea that multinationals are such largely thanks to knowledge advantages—unless it is the case that these knowledge advantages simply result from more H. These different possibilities our empirical work will try to discern.

A related important issue is that each idea might not be equally important to all firms. Consider, for example, the stock of knowledge at an industry trade fair or at a university. Each firm visiting the trade fair or co-operating with a university may or may not gain ideas from such forums because not all ideas are of equal importance to all firms. Thus, as is well known we wish to measure not only the stock of knowledge but the weight in the knowledge flow that that particular source has for a given firm (just as citation-weighted patents have been used in some studies). One might think of (1) as imposing both that all ideas are available from all firms *and* that each idea carries equal weight for all firms. Thus, we may substitute out A in (1) terms of the knowledge stock inside and outside the firm (A_i and $A_{\_i}$) times the flow of ideas from those stocks (θ_{1i} and θ_{2i}):

$$\Delta A_i = f(\theta_{1i} A_i, \theta_{2i} A_{\_i}, H_i, MNE_i) \qquad (2)$$

where we have generalised the functional form and we have also added indicator variables for various dimensions of global engagement such as being part of a multinational enterprise (*MNE*).

Equation (2) presents us with several alternative specifications for knowledge production functions. For example, estimating (2) with just the *MNE* indicators (plus any other appropriate controls not in (2)—e.g., industry dummies, see below) summarizes whether globally engaged firms generate more knowledge output than do purely domestic firms. As we have discussed, however, greater ΔA for globally engaged firms might just reflect their greater use of knowledge inputs. Thus, estimating (2) with the *MNE* indicators plus H would examine the hypothesis that globally engaged firms enjoy different (presumably larger) stocks of knowledge. Thanks to the richness of our CIS data, we can then add direct measures of sources of knowledge $\theta_{1i}A_i$ and $\theta_{2i}A_{\_i}$ to try to explain away any significant role captured by the *MNE* indicators.

Estimating the Knowledge Production Function

A number of issues arise in estimating equation (2). One overarching issue is how best to measure the regressand and regressors. Our data present many possible measures of ΔA: firms report whether they have had a process or product innovation; whether such innovations are novel to their industry; what fraction of sales are accounted for by the innovations; and also the number of patents filed. There are pros and cons to different measures: patents are continuous, but patenting is an endogenous firm choice conditional on innovating that depends on a host of considerations such as legal systems and internal labor markets. For reasons including breadth, our benchmark analysis will use for ΔA the dichotomous indictor of whether the firm had a process or product innovation. Clearly, choice of ΔA influences choice of estimator. Binary measures such as our benchmark regressand we estimate by probit. In contrast, when we measure ΔA as the fraction of sales due to the new innovation, we have a regressand covering the [0,1] interval with a large mass of firms at 0. In this case we estimate (2) by tobit.[1]

To measure knowledge inputs H, there may be more to the innovation process then just a head count of scientists. Scientists might vary by quality and the tools they have to work with (e.g., laboratories); and innovative ideas might come from other types of employees (e.g., management consultants or production-line workers who discover process innovations). Our benchmark analysis will use for H employment data on scientists, but we experiment with several alternatives including adding other educated workers and also measuring the financial expenditures on these innovative workers and/or capital to work with them.

[1] There are also firms who report zero for any inputs and zero for ΔA (almost no firms report zero for any inputs and $\Delta A>0$).. The data therefore consists of firms who have all zeros, i.e., those who did not try innovating at all, and those who have non-zero inputs but $\Delta A \geq 0$. Relative to the double-hurdle model, tobit imposes the same impact of inputs on the two types of firms. However, we think this is a reasonable assumption, since, absent unobservables, presumably if scientists were free firms who do not hire them now would hire them.

To measure stocks of knowledge and the related flow of ideas, it is important to emphasize that flows from these stocks to firms may or may not be mediated by market transactions. Market-mediated knowledge flows would include joint research projects with other institutions such as universities. Non-market-mediated knowledge flows, the knowledge "spillovers" or externalities discussed in the introduction, would include things like serendipitous conversations with academics at a trade fair. Ideally our data would cleanly distinguish these alternative modes of accessing stocks of knowledge. In practice this will not be the case.

A second important estimation issue to consider is endogeneity. Inputs to knowledge creation H are chosen by firms; if these choices are correlated with unobservables in equation (2), then coefficient estimates in (2) might be biased. Methods for addressing possible endogeneity bias in production-function estimates have been developed in recent years in the IO literature (see Haskel, et al 2002 for a discussion). The general approach of these methods is to proxy for unobserved determinants of input choices with observables: e.g., to proxy for unobserved managerial talent using firm fixed effects or firm-and-time-varying levels and exponentials of observable inputs such as investment or materials.

For our purposes, in principle we could create a panel and introduce fixed effects by using the previous CIS survey wave (covering 1994-1996). This approach would present its own problems, however: there were substantive differences in question wording across the two waves; there is survivor bias in such a panel; and probit and tobit estimators with fixed effects tend to suffer from unknown biases (Greene, 2004a and 2004b). For a more detailed discussion see Criscuolo and Haskel (2003). The obvious alternative is to instrument for input choice H. Exploiting the regional variation in our data, we experimented with using data on local wages of scientists as a possible instrument. But we found local wages and scientist employment to be very weakly correlated (not surprising if the relevant labor market for U.K. scientists is largely national, in which case observed wage variation is largely spurious), which means any IV estimates would likely suffer from bias due to weak instruments (Staiger and Stock, 1997).

Given these endogeneity concerns, one might prefer to interpret our estimates below as a way to uncover patterns than can be closely interpreted by theory. Even with this restrictive interpretation, we think our results are of interest thanks both to their close link to the underlying macro/IO theory and to the richness of our data.

3. Data Description

Our empirical analysis uses a cross-section data set of U.K. firms constructed from three key data sources. First is the U.K. Community Innovations Survey. This is an EU-wide survey, administered by the OECD, developed to measure both innovate output and inputs of firms. The other two data sources are the Annual Inquiry into Foreign Direct Investment (AFDI) and the Annual Respondents Database (ARD), both of which are used to tag each of our U.K. CIS observations as a parent of a U.K.-based multinational or an affiliate of a foreign-owned multinational. We briefly discuss each data source.

CIS Data

The U.K. CIS is part of an EU-wide survey that asks companies to report the output of their innovation efforts (introduction of innovative new products and/or processes; percentage of sales arising from new and improved products; and "soft" innovations, such as organisational change); the firm's inputs to innovations (R&D, scientists, etc.) and the sources of knowledge for innovation efforts. There have been three waves of U.K. CIS surveys: CIS1 (covering the period 1991-3), CIS2 (1994-6) and CIS3 (1998-2000). Our work covers CIS3.

The CIS is a voluntary postal survey carried out by the Office of National Statistics (ONS) on behalf of the Department of Trade and Industry (DTI). ONS selects survey recipients by creating a stratified sample of firms with more than 10 employees drawn from the Inter-Departmental Business Register (IDBR) by SIC92 two-digit classes and eight employment-size bands. The survey covers both the production and the service sectors.[2] CIS3 was in the field twice: the first wave sampled 13,340 enterprises, and the second top-up covered 6,285 to make the sample representative at the regional level. Of the total 19,625 enterprises to which the survey was sent, 8,172 responded (3,605 in services and 4,567 in production), for an overall response rate of 42%.

Two important issues immediately arise from the sample design of the CIS. First is the question of non-response. Since the survey is voluntary and postal, there is the risk of low-response and thus of non-response bias.[3] For CIS3 we investigated the characteristics of the firms that returned the questionnaire relative to non-respondents using the matched CIS survey universe (i.e., the firms chosen to be surveyed regardless of whether they reply or not). Non-respondents were on average larger than respondents, both in terms of turnover and employment.

Second, the survey was conducted at the enterprise level; where enterprise is defined as "the smallest combinations of legal units which have a certain degree of autonomy within an enterprise group." Thus, an enterprise is roughly a firm, where each firm can have more than one business establishment and can also be part of a larger multi-enterprise business entity called an enterprise group. For our interest in globally engaged firms, by construction any U.K. enterprise that is part of a multinational firm has at least one other enterprise in its enterprise group somewhere in the world. One might worry about reporting error due to respondents not

---

[2] Production includes manufacturing; mining; electricity, gas and water; and construction. Services includes wholesale trade; transport, storage, and communication; and financial intermediation and real estate.

[3] To boost response enterprises were sent the survey, posted a reminder, posted a second reminder (with the survey again) and finally telephoned.

answering at the desired enterprise level. We were able to identify small numbers of such probable cases through data checking and cleaning; our results appear to be robust to this issue.[4]

Having explained the structure of collecting the CIS data, let us briefly summarize its contents. The essential idea of the survey is to get enterprises to retrospectively report separately "technological" and "organisational" innovations (with separate questions for each group). In turn, technological innovations are split into process and product innovations. Table 1 sets out the survey questions for both product and process innovation (as well as for many other key variables in our analysis). For both these (and many other) key variables, the CIS questionnaire provided additional information (e.g., example responses) to clarify definitions and thereby improve response quality. For each firm we constructed our baseline measure of innovation, *Innovation*, as a dichotomous variable equal to one if the firm reported yes to either (or both) process and product innovation.

A potential problem with this approach is its subjective nature. Since companies are asked about products or processes that are "technologically new," there is obvious scope for differences in interpretation even with guiding information provided in the questionnaire. For example, some firms might report "yes" to all questions to not be seen as somehow backward. All this may introduce measurement error into our regressand in equation (2).[5]

Our possible measures of *H* and *A* from equation (2) also appear in Table 1. For H we can use variables such as *Proportion Engineers*. To measure A we use various combinations of the component responses to *Information Sources*. For each of these possible categories of information, firms were asked to report whether any information from this source, and, if so,

[4] These robustness checks consist, for example, in leaving in the sample only single-plant firms. Indeed, a misunderstanding can arise only for multi-plant firms and/or firms that are part of an enterprise group.

[5] One check on these responses is that companies are asked to fill out in longhand their "most important product or process". The long-hand response rates were not high (about 30%) but our casual inspection of these responses relative to the guidelines provided indicated that enterprises were able to report technological innovations.

whether its importance was low, medium, or high. Following the precedent of other researchers who have used CIS-type data for other EU countries (e.g., Cassiman and Veuglers, 2002), we translated these qualitative responses into a categorical variable with values 0, 1/3, 2/3, and 1 going from no information to information of high importance.[6] Again, we consider *Information Source*s to be an especially valuable piece of the overall data analysis because it allows us to test directly the idea that different firms—perhaps along the lines of their global engagement—have access to different knowledge stocks.[7]

AFDI and ARD Data

The CIS measures one dimension of the global engagement of firms in that it asks whether and how much firms exported in 1998 and again 2000. But it does not have any information on the other key dimension of global engagement: being part of a multinational firm. Accordingly, to add this information to the CIS data we merged in nationality of ownership data from the AFDI and ARD.

The AFDI is an annual survey to businesses which requests a detailed breakdown of the financial flows between UK enterprises and their overseas parents or subsidiaries. It contains an "outward" part that measures outward FDI by U.K. parents and also an "inward" part that measures FDI into the U.K. by foreign-owned firms. To run the AFDI, ONS maintains a register that holds information on the country of ownership of each enterprise and on which U.K. enterprise has foreign subsidiaries or branches. This register is designed to capture the universe of enterprises that are involved in FDI abroad and in the UK, where a 10% ownership stake is applied in both directions. It is continuously updated from a variety of government and private-sector data sources.

[6] To construct this variable, we used the maximum value of the response to the combination of questions divided by three.
[7] The entire CIS3 survey can be seen at http://www.dti.gov.uk/iese/cis_quest.pdf.

The ARD provides an alternative source of information on the country of ownership of foreign-owned firms operating in the U.K., where here the underlying data source is solely Dun & Bradstreet Global "Who Owns Whom" database. The AFDI and ARD methods differ in two potentially important respects: AFDI tracks the nationality of *direct* owners using a *threshold* of 10%; ARD tracks the nationality of *ultimate* owners using a *threshold* of 50%. In principle, then, these two different data sets can yield different answers as to whether a U.K. firm is foreign owned, and, if so, by a firm in what country. In practice, there were very few such discrepancies in our data: only about two dozen firms classified as foreign owned by AFDI but not by ARD. We chose to use the AFDI categorization in these "conflicting cases," both to maximize the number of foreign-owned observations and because its 10% ownership criterion is widely used by statistical agencies in other countries (e.g., the United States' Bureau of Economic Analysis). In practice, our results are robust to giving precedence to the ARD scheme.

We were able to merge accurately the AFDI and ARD data into the CIS data since the ONS used the same core set of firm and group identifiers for all three data sets. With all this information combined, we created four categories of global engagement for our firms: parents of U.K. multinationals; affiliate of foreign multinationals; non-multinational firms that export; and "purely domestic" firms that neither export nor are part of a multinational.[8]

For our overall sample of 8,172 firms we ended up with the following distribution: 636 multinational parents (7.78% of the sample); 735 multinational affiliates (8.99%); 1,927 non-multinational exporters (23.58%); and 4,874 purely-domestic firms (59.64%). Consistent with many of the studies cited in the introduction, in our sample there are basic performance

[8] There were also a small number of firms that were classified as U.K. parents in the AFDI data but also U.K. affiliates in the ARD data. In principle, such complicated ownership structures can be found given the nature of the two data sets. In practice, to maximize our number of U.K. parents we placed this small number of firms in the U.K.-parent category. Our results below were robust to the alternative of placing them in the U.K.-affiliate group.

differences across these four groups. For example, mean firm size (either sales or employment) and capital intensity are highest for the parents, then the affiliates, then the exporters, and finally the purely domestics. The same ordering also applies for the fraction of firms that have more than one establishment within the U.K. We now turn to our data analysis of the innovative activity of these different firms.

4. Data Analysis and Estimation Results

Table 2 presents summary statistics on innovation for our entire sample of firms and also for our four sub-samples by global engagement. There are two important messages from Table 2. First, globally engaged firms create substantially more new ideas than do purely domestic firms. For our broad benchmark measure of making new ideas, *Innovate*, about 45% of all multinational firms and 40% of all exporters report having innovated. In contrast, only 21.2% of purely domestic firms report innovating. A similar contrast appears for alternative measures of knowledge output. Column 2 shows a similar pattern for *Patent Protect*, a binary variable equal to one if the firm reports either having applied for new patents 1998-2000 or using existing patents to protect its innovations. Column 3 again shows a similar pattern for the number of new patents applied for, *Patents* (*Patent Protect* is somewhat broader than *Patents* since a firm might rely on patent protection even if it did not recently apply for new patents). That said, it is important to note that within all sub-samples the distribution of knowledge creation seems highly skewed, with the median firm reporting no innovative activity. Despite this intra-group skewness, this overall picture of globally engaged firms making more new ideas holds across all our other possible measures of knowledge creation (e.g., the various components "within" *Innovate*). This is our first key finding of the paper.

The second important message of Table 2 is that globally engaged firms use more inputs for making new ideas. All three categories of globally engaged firms allocate around 10% of their total employment to knowledge-creating occupations such as scientists and engineers, as indicated by *% Scientists*. This share is more than double the 4.1% in the average purely-domestic firm. The production-function framework motivating our analysis suggests that some—or perhaps all?—of the variation in knowledge outputs in the first three columns of Table 2 can be accounted for by this variation in knowledge inputs in the final column. To answer this systematically, we now estimate versions of the ideas production function in equation (2).

Table 3 reports regression results for several specifications of the ideas production function, all of which use as the regressand our baseline knowledge-output measure *Innovate.* Each column of Table 3 corresponds to a different specification estimated via probit, with each row in a column reporting the marginal coefficient estimate (and robust standard errors, clustered on firm within an enterprise group) for the indicated regressor. All specifications in Table 3 include a common set of control regressors (not reported for brevity) to help control for plausibly important cross-firm sources of innovative heterogeneity: a full set of two-digit industry dummies (approximately 50 controls), a full set of regional dummies (13 controls); firm size (measured as total employment), and a categorical indicator for various responses to the question *Structural Changes* in Table 1 (to account for patterns such as newly-born start-up firms being more likely to innovate a lot).

Column 1 estimates an ideas production function as in equation (1), where the maintained assumption is that all firms have equal access to the same stock of knowledge (and thus that the influence of this stock is subsumed into the constant term). As expected, there is a strongly positive correlation between researchers employed and the likelihood of creating knowledge.

Column 2 adds to the production function in Column 1 our various indicators of global engagement, where our omitted group is the purely domestic firms. Thus, we are now estimating a simple version of equation (2). If the maintained hypothesis of a seamless global stock of knowledge equally accessible to all firms were true, then these global-engagement indicators should be individually and jointly insignificantly different from zero. The coefficient estimates strongly reject this hypothesis. Instead, all three indicators are statistically and economically large. Exporting firms are about 12 percentage points more likely to report innovative output relative to purely domestic firms; affiliates of foreign-owned multinationals are about 14 percentage points more likely; and parents of U.K.-owned multinationals are about 17 percentage points more likely.

Column 3 is consistent with different firms having access to systematically different stocks of knowledge based on their global engagement. It suggests that the fact from Table 2 that globally engaged firms generate more knowledge output is not just because these firms use more knowledge inputs. It is also because of the knowledge stocks to which these firms have access. This finding accords with the discussion earlier in the paper that multinationals and exporters look like "better firms" in terms of knowledge activity.

Without the richness of our CIS data, at this point our analysis might have to conclude with speculation about what exactly are the sources of greater information to which globally engaged firms seem to have access and thus greater knowledge production. With these data, however, we attempt to unbundle what the global-engagement dummies in Column 2 represent by using all the self reports of where firms learn from in the *Information Sources* questions.

Columns 3 through 11 of Table 3 do precisely this. Column 3 adds in self reports of learning from the overall enterprise group. For any firm in the U.K. that is part of a multinational, we might expect that at least some of its differential learning comes from the firm-wide knowledge its overall enterprise group creates elsewhere in the world. The dramatic drop in size and significance on the two multinational indicators suggests that not just some but rather virtually all of the differential innovative output of U.K. firms that are part of multinationals is because of their access to intra-firm knowledge.

The size of the coefficient estimate on exporting firms also drops, which suggests that like multinationals they too learn from other operations within their enterprise group in the U.K. But both because this estimate remains significantly positive, and also because the enterprise-group regressor might be correlated with other self-reported sources of knowledge, in columns 4 through 11 we expand our specification of the production function by adding in additional sources. In columns 4 through 10, adding additional information sources one at a time does not materially change the message of column 3, with the important exception in column 4 that vertical information—i.e., from customers and suppliers—appears to be about as important for globally engaged firms as does their intra-firm worldwide information. Adding all information sources together as in the final column 11 reveals, not surprisingly, substantial collinearity among many of the information sources. As in column 4, what stands out is how globally engaged firms appear to create more output thanks to their greater access to knowledge both from other parts of their firm and also from links with customers and suppliers.

Table 4 replicates the analysis of Table 3, but now for a different measure of innovation output, *Patent Protect* as defined in Table 2. Reading across the columns tells the same broad story as that of Table 3: knowledge output varies strongly with knowledge inputs; globally engaged firms generate more knowledge output, conditional on their use of knowledge inputs;

much (though here definitely not all as in Table 3) of this global-engagement effect appears driven by these firms having access to larger stocks of knowledge.

The most notable difference across the two tables is that the global engagement indicators remain much more important in Table 4. This is an area for future investigation; one possibility discussed earlier in the paper might be that patenting is a noisier measure of pure knowledge output because it entails not just creating knowledge but the endogenous choice to actively seek protections for that knowledge, presumably from competitors. Consistent with this possibility, notice that in Table 4 information from competitors remains a much stronger regressor in the final column 11.

The results in Tables 3 and 4 are robust to a number of measurement and specification choices. In particular, the general impacts of our global-engagement and information-source regressors do not change when we vary the exact H measure of firm use of innovation-producing workers. Results also do not change when we vary the set of control regressors: e.g., using firm sales instead of firm employment for size, or dropping either the industry or region dummies. We also estimated a large number of specifications interacting our global-engagement indicators with other regressors—e.g., with our measures of H to see if globally engaged firms enjoy higher marginal productivity from knowledge inputs. These interactions almost always were insignificantly different from zero, which reinforces our interpretation above that globally engaged firms create more new knowledge in part because they have access to larger stocks of existing knowledge.

5. Conclusions

In this paper we have tried to better understand creating new ideas. Our approach has been to estimate knowledge production functions on a data set of thousands of U.K. firms for which we have detailed information on knowledge outputs, inputs, and—importantly—flows from various knowledge stocks. We focused in particular on the hypothesis from the trade literature that globally engaged firms—either multinationals or exporters—have access to larger knowledge stocks. We found that globally engaged firms do generate more ideas than their purely domestic counterparts. This is not just because they use more knowledge inputs. Importantly, it is also because they have access to a larger stock of ideas through two main sources: their upstream and downstream contacts with suppliers and customers, and, for multinationals, their intra-firm worldwide pool of information.

These results offer an initial step in the direction recommended by Jones (2003) in his forthcoming literature survey of economic growth: examining real-world data. As such, the results suggest that future modelling needs to think harder about possible pitfalls from assuming a seamless global stock of knowledge to which all actors have easy and equal access. These results also inform the growing literature in trade on multinational firms. For example, the now-standard knowledge-capital model of multinationals is largely silent on how these firms optimally structure intra-firm knowledge sharing. In future work, we aim to apply our data to issues such as these.

References

Aitken, Brian J., and Ann E. Harrison. 1999. "Do Domestic Firms Benefit from Foreign Direct Investment? Evidence from Venezuela." *American Economic Review*, June, pp. 605-618.

Bernard, Andrew B., Michael M. Knetter, and Matthew J. Slaughter. 2004. *Global Economics for Managers*. New York: Addison-Wesley Publishers, forthcoming.

Brainard, Lael. 1997. "An Empirical Assessment of the Proximity-Concentration Tradeoff between Multinational Sales and Trade." *American Economic Review* 87: 520-544.

Carr, David L., James R. Markusen, and Keith E. Maskus. 2001. "Estimating the Knowledge-Capital Model of the Multinational Enterprise." *American Economic Review* 91, pp. 693-708.

Cassiman, Bruno, and Reinhilde Veugelers. 2002. "R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium." *American Economic Review* 92 (4), pp. 1169-1184.

Criscuolo, Chiara, and Jonathan E. Haskel. 2003. "Innovations and Productivity Growth in the UK: Evidence from CIS2 and CIS3." Draft paper, available at www.ceriba.org.uk.

Criscuolo, Chiara and Ralf Martin. 2003. "Multinationals and U.S. Productivity Leadership: Evidence from Great Britain." CeRiBa Discussion Paper.

Doms, Mark E., and J. Bradford Jensen. 1998. "Comparing Wages, Skills, and Productivity Between Domestically and Foreign-Owned Manufacturing Establishments in the United States." In R. Baldwin, R. Lipsey, and J. D. Richardson (eds.), *Geography and Ownership as Bases for Economic Accounting*. Chicago: University of Chicago Press, pp. 235-255.

Greene, William. 2004a. "Fixed Effects and Bias Due to the Incidental Parameters Problem in the Tobit Model." *Econometric Reviews*, forthcoming.

Greene, William. 2004b. "The Bias of the Fixed Effects Estimator in Nonlinear Models." *The Econometrics Journal*, forthcoming.

Griliches, Zvi. 1998. *R&D and Productivity: The Econometric Evidence*. Chicago: University of Chicago Press.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of R&D to Productivity Growth." *Bell Journal of Economics* (10), pp. 92-116.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1), pp. 83-116.

Haskel, Jonathan E., Sonia Pereira, and Matthew J. Slaughter. 2002. "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?" NBER Working Paper No.8724.

Jones, Charles I. 2003. "Growth and Ideas." Forthcoming in *Handbook of Economic Growth*.

Jones, Charles I. 2002. "Sources of U.S. Economic Growth in a World of Ideas." *American Economic Review* 92(1), pp. 220-229.

Keller, Wolfgang. 2001. "The Geography and Channels of Diffusion at the World's Technology Frontier." National Bureau of Economic Research Working Paper No. 8150, March.

Mansfield, Edwin and Anthony Romeo. 1980. "Technology Transfer to Overseas Subsidiaries by U.S.-Based Firms." *Quarterly Journal of Economics*, 95 (4), pp. 737-750.

Markusen, James R. 2002. *Multinational Firms and the Theory of International Trade*. Cambridge, MA: MIT Press.

Moran, Theodore H. 2001. *Parental Supervision: The New Paradigm for Foreign Direct Investment and Development*. Washington, D.C.: Institute for International Economics.

Parente, Stephen L., and Edward C. Prescott. 1994. "Barriers to Technology Adoption and Development." Journal of Political Economy 102 (2), pp. 298-321.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5), pp. S71-S102.

Staiger, Doug, and James Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65(3), pp. 557-586.

Tybout, James R. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, 38, pp. 11-44.

Table 1: Variable Definitions in CIS3

1. Measures of Innovation Outputs

Variable Name	***Variable Definition***
Number of Patents	How many patents, if any, did your enterprise apply for during the period 1998 to 2000?
Innovation expenditures/ activity	**Did your enterprise engage in the following innovation activities in 2000? Please estimate innovative expenditure in 2000, including personnel and related investment expenditures (no depreciation)**
Process Innovation	During the three year period 1998-2000, did your enterprise introduce any technologically new or improved processes for producing or supplying products which were new to your firm?
Novel Process Innovation	During the three year period 1998-2000, did your enterprise introduce any new or significantly Improved processes for producing or supplying products (goods or services) which were new to your industry?
Product Innovation	During the three year period 1998-2000, did your enterprise introduce any technologically new or significantly improved products (goods or services) which were new to your firm?
% turnover due to new and improved products	Please estimate how your turnover in 2000 was distributed between products (goods or services) Introduced during the period 1998-2000 which were: New to your firm + Significantly improved (%)
Novel Product Innovation	During the three year period 1998-2000, did your enterprise introduce any new or significantly Improved products (goods or services) which were also new to your enterprise's market?
% turnover due to novel product innovation	Please estimate the share of turnover of these (novel) products in 2000

2. Measures of Innovation Inputs (H_i)

Innovation expenditures/ activity	**Did your enterprise engage in the following innovation activities in 2000? Please estimate innovative expenditure in 2000, including personnel and related investment expenditures (no depreciation)**
R&D Spending	Reported Intramural R&D + Acquisition of external R&D
Innovation Expenditure	Total expenditure when reported or the sum of reported Intramural R&D + Acquisition of external R&D + Acquisition of machinery and equipment + Acquisition of other external knowledge + All design functions + Internal or external training + Internal or external marketing
Intramural R&D	_Please tick if expenditure in the category Internal R&D -If you have internal R&D-activities: How many persons were involved in R&D activities within your enterprise in 2000? (in full time equivalents) positive number How did your enterprise engage in R&D during the three year period 1998-2000? Continuously
Proportion Engineers	Approximate proportion educated to degree level or above Science and engineering subject (%)
Proportion Other Graduates	Approximate proportion educated to degree level or above Other subjects (%)

Table 1: Variable Definitions in CIS3 (Continued)

3. Measures of Information Flows

Part of a group	
	A range of factors may inhibit your ability to innovate. Please grade the importance of the Following constraints during the period 1998-2000:
Information Sources	**Please indicate the sources of knowledge or information used in your technological innovation activities, and their importance during the period 1998-2000.**
Enterprise-Group Information	Within the enterprise or other enterprises within the enterprise group
Vertical Information from Customers & Suppliers	Suppliers of equipment, materials, components or software + Clients or customers
Information from Competitors	Competitors
Commercial Information	Consultants+ Commercial laboratories/ R&D enterprises
Free Information	Professional conferences, meetings+ Trade associations+ Technical/trade press, computer databases+ Fairs, exhibitions
Regulatory Information	Technical standards+ Environmental standards and regulations+ Health and safety standards and regulations
Information from Universities	Universities or other higher education institutes+ Private research institutes
Information from Government	Government research organisations+ Other public sector e.g. business links, Government Offices
Soft Innovation	Did your enterprise make major changes in the following areas of business structure and Practices during the period 1998-2000 and how far did business performance improve as a result? Implementation of :new or significantly changed Corporate strategies or advanced management techniques or Organisational structures

4. Other Variables

Employment	**Total employees at the enterprise**
Exports	**Export value at the enterprise, measure is 1/0 dummy if exports>0**
Structural Changes	**Did any of the following significant changes occur to your enterprise during the three year period 1998-2000?**
Just established	The enterprise was established
Merger	Turnover increased by at least 10% due to merger with another enterprise or part of it
Sale or closure of part of the enterprise	Turnover decreased by at least 10% due to sale or closure of part of the enterprise

Table 2: Summary Statistics on Innovative Activity

Sub-Sample	Innovate	Patent Protect	Patents	% Scientists
Multinational Parents	0.455	0.318	9.282	0.122
(N = 635)	(0.000)	(0.000)	(0.000)	(0.050)
Multinational Affiliates	0.450	0.364	2.553	0.100
(N = 734)	(0.000)	(0.000)	(0.000)	(0.050)
Non-Multinational Exporters	0.401	0.229	0.829	0.086
(N = 1,925)	(0.000)	(0.000)	(0.000)	(0.030)
Purely Domestics	0.212	0.056	0.101	0.041
(N = 4,864)	(0.000)	(0.000)	(0.000)	(0.010)
All Firms	0.297	0.148	1.276	0.063
(N = 8,158)	(0.000)	(0.000)	(0.000)	(0.020)

Notes: For each cell, indicated summary statistics are means (and medians in parentheses). *Innovate* is an indicator variable equal to one if firms reported any process or product innovation. *Patent Protect* is an indicator variable equal to one if firms reported either having applied for new patents 1998-2000 or using existing patents to protect its innovations. *Patents* is the number of patents applied for in the last three years. *% Scientists* is the share of firm employment accounted for by scientists, engineers, and similar occupations. See text for data details.

Table 3:
Estimates of the Knowledge Production Function
For Output Measure *Innovate*
(1/0 variable equal to one if firms reported any process or product innovation)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
% Scientists	0.3696	0.3644	0.1288	0.1339	0.1338	0.1235	0.1243	0.1270	0.1150	0.1286	0.1225
	(0.0440)***	(0.0498)***	(0.0510)**	(0.0496)***	(0.0500)***	(0.0508)**	(0.0495)**	(0.0509)**	(0.0511)**	(0.0507)**	(0.0496)**
Exporter		0.1227	0.0393	0.0418	0.0382	0.0383	0.0400	0.0430	0.0369	0.0411	0.0385
		(0.0169)**	(0.0174)**	(0.0173)**	(0.0174)**	(0.0174)**	(0.0173)**	(0.0174)**	(0.0174)**	(0.0174)**	(0.0172)**
Multinational Parent		0.1666	0.0169	0.0220	0.0112	0.0074	0.0133	0.0159	0.0063	0.0147	0.0176
		(0.0292)**	(0.0276)	(0.0276)	(0.0275)	(0.0271)	(0.0272)	(0.0274)	(0.0270)	(0.0275)	(0.0272)
Multinational Affiliate		0.1350	-0.0002	0.0245	-0.0031	-0.0073	0.0089	0.0007	-0.0090	-0.0004	0.0246
		(0.0263)**	(0.0243)	(0.0252)	(0.0241)	(0.0240)	(0.0245)	(0.0243)	(0.0239)	(0.0243)	(0.0252)
Enterprise-Group Info			0.6738	0.4261	0.6251	0.6210	0.5429	0.6198	0.6483	0.6511	0.4145
			(0.0184)***	(0.0219)***	(0.0202)***	(0.0198)***	(0.0208)***	(0.0216)***	(0.0190)***	(0.0190)***	(0.0230)***
Government Info										0.1444	-0.0021
										(0.0347)***	(0.0374)
University Info									0.1496		0.0526
									(0.0327)***		(0.0349)
Gov't Regulations Info								0.1052			-0.0819
								(0.0224)***			(0.0253)***
Free Info							0.2722				0.1516
							(0.0237)***				(0.0285)***
Commercial Info						0.1672					0.0489
						(0.0266)***					(0.0282)*
Competitors' Info					0.1353						-0.0927
					(0.0257)***						(0.0284)***
Suppliers/Customers Info				0.3727							0.3579
				(0.0218)***							(0.0272)***
# Observations	6557	5665	5570	5570	5570	5570	5570	5570	5570	5570	5570

Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4:
Estimates of the Knowledge Production Function
For Output Measure *Patent Protect*
(1/0 variable equal to one if firms reported any use of patent protection)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
% Scientists	0.2685	0.1864	0.1027	0.1027	0.1075	0.0937	0.1002	0.1010	0.0825	0.1024	0.0883
	(0.0290)***	(0.0315)***	(0.0289)***	(0.0288)***	(0.0279)***	(0.0285)***	(0.0282)***	(0.0282)***	(0.0296)***	(0.0281)***	(0.0282)***
Exporter		0.1321	0.0870	0.0863	0.0848	0.0855	0.0851	0.0898	0.0835	0.0892	0.0845
		(0.0144)***	(0.0131)***	(0.0131)***	(0.0129)***	(0.0130)***	(0.0130)***	(0.0132)***	(0.0131)***	(0.0131)***	(0.0129)***
Multinational Parent		0.2567	0.1663	0.1664	0.1595	0.1552	0.1631	0.1661	0.1463	0.1640	0.1418
		(0.0300)***	(0.0256)***	(0.0257)***	(0.0253)***	(0.0253)***	(0.0258)***	(0.0256)***	(0.0247)***	(0.0256)***	(0.0246)***
Multinational Affiliate		0.2953	0.2142	0.2170	0.2083	0.2044	0.2155	0.2144	0.1980	0.2138	0.1919
		(0.0270)***	(0.0256)***	(0.0257)***	(0.0255)***	(0.0254)***	(0.0258)***	(0.0257)***	(0.0257)***	(0.0257)***	(0.0256)***
Enterprise-Group Info			0.1906	0.1590	0.1449	0.1466	0.1443	0.1460	0.1515	0.1644	0.1161
			(0.0109)***	(0.0142)***	(0.0115)***	(0.0113)***	(0.0122)***	(0.0122)***	(0.0110)***	(0.0109)***	(0.0142)***
Government Info										0.1374	0.0263
										(0.0183)***	(0.0206)
University Info									0.1755		0.1160
									(0.0181)***		(0.0193)***
Gov't Regulations Info								0.0814			0.0270
								(0.0129)***			(0.0147)*
Free Info							0.0930				0.0197
							(0.0138)***				(0.0170)
Commercial Info						0.1223					0.0562
						(0.0147)***					(0.0158)***
Competitors' Info					0.1136						0.0787
					(0.0144)***						(0.0166)***
Suppliers/Customers Info				0.0470							-0.0487
				(0.0144)***							(0.0180)***
# Observations	6129	5293	5235	5235	5235	5235	5235	5235	5235	5235	5235

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%