

TRADE AND R&D SPILLOVER
EFFECTS: IMPLICATIONS FOR
FIRM LEVEL ANALYSIS IN THE
AGRICULTURAL SECTOR

by

David Skully and Manitra A. Rakotoarisoa

Food and Agriculture Organization of the United Nations



The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO) concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO in preference to others of a similar nature that are not mentioned.

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO.

© FAO 2013

FAO encourages the use, reproduction and dissemination of material in this information product. Except where otherwise indicated, material may be copied, downloaded and printed for private study, research and teaching purposes, or for use in non-commercial products or services, provided that appropriate acknowledgement of FAO as the source and copyright holder is given and that FAO's endorsement of users' views, products or services is not implied in any way.

All requests for translation and adaptation rights, and for resale and other commercial use rights should be made via www.fao.org/contact-us/licence-request or addressed to copyright@fao.org.

FAO information products are available on the FAO website (www.fao.org/publications) and can be purchased through publications-sales@fao.org.

CONTENTS

INTRODUCTION	1
PART I: R&D SPILLOVERS: MACRO-LEVEL ANALYSIS	2
Introduction	2
TFP – total factor productivity and growth theory	3
Knowledge, innovation and R&D spillovers	5
Endogenous growth theory	6
Measuring R&D and R&D spillovers	9
Conclusion Part I	12
PART II: R&D SPILLOVERS: MICRO-LEVEL ANALYSIS	14
Introduction	14
Micro-level analysis of productivity	15
Heterogeneous firm dynamics: empirical methods	17
Heterogeneous trading firm dynamics	20
Selection, learning and innovation at the industrial and national level	21
Heterogeneous trading firms and related-party trade	26
Conclusion Part II	27
PART III: CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH AND AGRICULTURAL TRADE	28
APPENDIX A: MEASURING STOCKS OF KNOWLEDGE	31
APPENDIX B: GROSSMAN AND HELPMAN ON WEIGHTING INTERNATIONAL INFORMATION FLOWS	32
BIBLIOGRAPHY	33



ACKNOWLEDGEMENTS

This paper is part of the series of EST/FAO studies on “Analysis of the Size and Distribution of Trade Impacts on Agro-Trading Firms and Industries”, funded by the Government of Australia under the Project MTF/GLO/356/AUL. The authors thank Jamie Morrison and Ramesh Sharma for initiating the projects. Technical support from Rita DiIorio is gratefully acknowledged. Comments and suggestions can be sent to david.skully@gmail.com and manitra.rakotoarisoa@fao.org.

INTRODUCTION

This paper provides a critical review of the economic literature on the relationship between international trade and research-and-development (R&D) spillovers with the aim of finding ways to track the spillover effects of agricultural trade at firm and industry levels. The term R&D spillover refers to the spread or diffusion of knowledge between countries, industries, or firms. New technological and managerial information can be used by firms to increase efficiency, which leads to improved aggregate income and welfare. Innovation and the application of knowledge is the central engine of economic growth in both the growth theory of the 1950s and 1960s and in endogenous growth theory, which emerged in the late 1980s.

The first part of the paper focuses on the empirical macro-level analysis of international R&D spillovers, in which the country is the unit of analysis. The key relationship is between macro-level productivity and R&D, domestic and foreign. The central question is whether and how international trade causes or influences macro-level productivity growth through international R&D spillovers. The discussion opens with a description of the seminal empirical analysis, Coe and Helpman (1995). Subsequent sections provide the larger theoretical context (growth theory) and discuss the measurement of the key variables: total factor productivity, the stock of R&D capital and estimating spillovers. The first part of the paper concludes with an assessment of the research to date. The assessment is that although the empirical evidence is consistent with the existence of trade-related R&D spillovers hypothesized by endogenous growth theory, the macro-level data do not provide sufficient information to determine how R&D spillovers occur. Because of the limitations of macro-level data, in the last 10 years economists have turned to the construction of micro-level data and the application of micro-level statistical methods to investigate how trade influences productivity.

The second part of the paper turns to the micro-level analysis of productivity change where the individual firm is the unit of analysis. It focuses on heterogeneous-firm models of international trade. The starting point of this literature is the fact that firms differ: they have different productivity levels. This is contrary to the macro-analytical assumption that the economy is composed of multiple representative firms that are homogeneous in equilibrium. Firm-level analysis reveals that relatively few firms engage in international trade; and those that do tend to be larger and more productive than the average firm. Similarly, relatively few firms engage in R&D, and those that do are typically large and highly productive. A central empirical question is whether engaging in international trade causes firms to become more productive. The literature finds some evidence that engaging in trade leads to higher firm productivity through “learning effects” both post pre-trade and post-trade. But the preponderance of evidence is for “selection effects”: high-productivity firms choose or self-select to engage in international trade. Trade is correlated with productivity growth, but the causal chain is primarily through trade inducing heightened competition and market selection; that is, by less productive firms exiting the market and the reallocation of resources among surviving and entering firms. Thus, the potential contribution of international R&D spillovers is greatly diminished when the trade-productivity relationship is examined at the firm level.

The third part of the paper summarizes the review and discussion, and anticipates some implications for future research. The application of the theory underlying the R&D spillovers of international trade in the agriculture sector in developing countries has remained puzzling. This review explores ways to assess the R&D spillovers of trade at sector and firm levels and contributes to the analysis of the size and impact of agricultural trade at firm and industry levels in developing countries.

PART I: R&D SPILLOVERS: MACRO-LEVEL ANALYSIS

Introduction

The 1980s witnessed two related innovations in economic theory. First, new trade theory emerged at the beginning of the 1980s. It introduced concepts from industrial organization into traditional trade theory; imperfect or monopolistic competition and scale economies play a central role. Second, building on the insights of new trade theory, new or endogenous growth theory emerged toward the end of the 1980s. Monopolistic competition is also central to endogenous growth theory: it provides a coherent incentive for innovative activity, commonly measured as R&D (research and development). Grossman and Helpman's (1991) book, *Innovation and Growth in the Global Economy*, is a synthesis of new trade theory and new growth theory; it remains the canonical text, although subsequent empirical research has induced some changes in theory.

One of the first empirical tests of the many propositions derived by Grossman and Helpman (1991) is a paper by Coe and Helpman (1995), "International R&D Spillovers." The hypothesis is that technological knowledge, measured as R&D, developed in one country spills over to other countries. The spill-in of knowledge can result in increased productivity and growth in the recipient country. Coe and Helpman (1995) test two related hypotheses: first, whether such R&D spillovers exist and second, whether R&D spillovers are positively related to trade. They estimate the following equations for a cointegrated panel of 22 developed countries for the years 1971-1990.¹

$$[1] \quad TFP_i = \beta_i^0 + \beta_i^d(R\&D_{domestic}) + \beta_i^f(R\&D_{foreign}) + \varepsilon_i$$

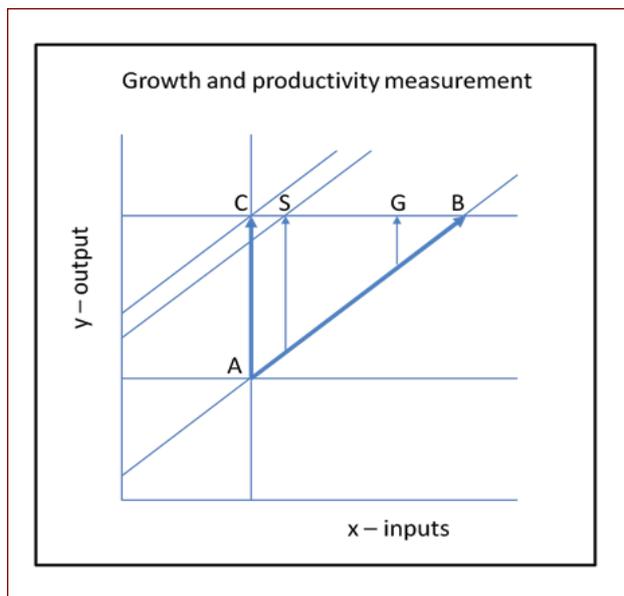
$$[2] \quad TFP_i = \beta_i^0 + \beta_i^d(R\&D_{domestic}) + \beta_i^m(m_i * R\&D_{foreign}) + \varepsilon_i$$

The first equation expresses a country's total factor productivity (in a given year, time subscripts are suppressed) as a function of the country's own (domestic) stock of R&D capital and the combined stocks of R&D capital of the other 21 (foreign) countries. The coefficients for these two variables are found to be positive and significant. The former is consistent with a country's own R&D contributing to its TFP growth; the latter is consistent with the existence of international R&D spillovers.

The second equation multiplies the foreign R&D term by m_i , the country's imports of goods and services as a proportion of its gross domestic product; this is a measure of a country's import intensity. The coefficients for both variables are found to be positive and significant. For the import-intensity-weighted foreign-R&D variable, this finding is consistent with the hypothesis that there is a positive relationship between trade, measured as import intensity, and international R&D spill-ins.

The next several sections are devoted to de-constructing Coe and Helpman (1995) and placing the paper in its larger theoretical context; in particular, explaining why R&D spillovers are important in endogenous growth theory. Also discussed are problems in measuring the variables in the equations above: total factor productivity, stocks of knowledge or R&D capital, and the empirical representation of R&D spillovers.

¹ The paper also includes other specifications. Note: variables in the estimation equations are in logarithms.



TFP - total factor productivity and growth theory

Total factor productivity (TFP) occupies a central place in economic growth theory. Productivity is a measure of the relative efficiency of a production system. The concept is simple: one calculates the ratio of output to input. A higher ratio (more output per input) indicates greater efficiency or higher productivity. Total factor productivity (sometimes called multi-factor productivity) measures the combined productivity of all factors of production: capital, labour, energy, and materials.²

The figure above illustrates the relationship between growth and TFP. It plots output (y) as a function of all inputs (x); everything is in per capita terms. Suppose point A represents an economy at time zero and that several years later its output increases to the level of points C, S, G, and B. There are numerous growth paths; consider the two extreme cases. The path from A to B represents growth solely through factor accumulation. Current savings are invested in capital and education. More inputs yield more output, but there is no TFP growth: the ratio of y to x is unchanged. The path from A to C is pure TFP growth. Inputs are constant, but output increases from A to C: all growth comes from the increase in the ratio of y to x; the economy combined inputs more efficiently. Solow (1957) analyzed the contributions of factor accumulation and TFP for the US economy from 1909 to 1949. He concluded that TFP growth accounted for seven-eighths of the observed growth in per capita output; factor accumulation accounted for one-eighth. This path would run from A to point S in the figure.

Solow, modeling the economy as if it were a single production unit, used a Cobb-Douglas production function restricted to have constant returns to scale (the sum of β_k and β_l equals one in equation [3]). Any growth in output not stemming from increases in capital or labour is attributed to changes in 'A', representing a change in the efficiency with which factors are combined. Solow referred to changes in 'A' as technological change and asserted that productivity growth is technological and finds its origin in innovative activity or R&D.

$$[3] \quad Y_t = A_t K_t^{\beta_k} L_t^{\beta_l}$$

$$[4] \quad y_t = \beta_0 + \beta_k k_t + \beta_l l_t + \varepsilon_t$$

$$[5] \quad tfp_t = \ln(A_t) = \beta_0 + \varepsilon_t = y_t - \beta_k k_t - \beta_l l_t$$

TFP growth is, in fact, a residual: it represents everything that is not measured as input growth (see equations [4 & 5]). If factors are not quality-adjusted or if they are not valued properly (e.g., if observed

² TFP is preferred to single-factor measures, such as labour productivity because single-factor productivity varies with the use of other inputs. For example, labour productivity may be raised by more intensive use of capital.

factor prices do not equal marginal products) then TFP measurement will be biased. In the Cobb-Douglas construction, the benefits of increased ‘A’ are distributed proportionately across all factors in the economy. This is not how technological change manifests itself in the economy. Solow (1960, 1962) proposed the concept of ‘vintage capital’ to represent the stylized fact that new model capital equipment is more productive than earlier vintages. But heterogeneous capital proved difficult to reconcile with an aggregate production function; the concept went dormant and was revived in the 1980s in the micro-level research discussed in the second part of this paper.

There is a tension between accounting for productivity and explaining productivity. Growth accounting is intimately related to national income accounting; without national income accounting data it is impossible to measure macro-level productivity. Economists concerned with national income accounts are appropriately obsessed with proper measurement of inputs and outputs: accounts must balance: something does not come from nothing. A persistent problem is properly accounting for changes in the quality and variety of inputs and outputs. Solow’s (1957) estimates were not based on quality-adjusted input or outputs; Solow considered quality improvements to be a form of technical change.

Griliches (1963) identified several potential sources of error in aggregate TFP measurement. Equation [6] expresses TFP growth in terms of growth rates; s represents the factor share of capital, and $(1-s)$ the factor share of labour.³ The sources of error are: the quantity and quality of labour (l); the quantity and quality of capital services (k); the relative factor share (s) measure; unmeasured inputs; and economies of scale.⁴

$$[6] \quad tfp = y - sk - (1 - s)l$$

Jorgenson and Griliches (1969, 1972) acted on Griliches’ diagnosis. Adjusting inputs and outputs for changes in quality and factor-utilization and depreciation rates, they found that at least 70 percent of measured U.S. output growth could be attributed to factor accumulation. These adjustments shift the growth path rightward to point G in the figure. Jorgenson and Griliches emphasized that attributing growth to measured changes in input quality does not explain how or why the changes in quality occurred. But it clarifies the task of explaining growth because it distinguishes between measured contributions to growth and unmeasured contributions. They argued that only the latter, unmeasured (or immeasurable) elements belong in the productivity residual.

Debates over whether and how to adjust inputs for quality changes continue. For example, the information and communication technologies (ICT) sector is R&D intensive and exhibits large annual quality improvements. The average \$1000 2012 laptop computer is over 100,000 times more powerful than the mainframe computers of the early 1960s, which cost more than \$10-million 2012 dollars. If the exponential decline in computing costs is not factored into the measure of ICT capital, then ICT capital is understated and the productivity residual is erroneously increased. Despite efforts to harmonize national income accounting systems, OECD countries still differ in methods of quality adjustment for national income accounts. For the 1990s, the U.S. ICT deflator averaged -20 percent annually, the U.K. deflator was -13 percent and for Germany, -8 percent. ICT prices are global: they differ little between

³ That is, k is the growth rate of capital, etc. This equation assumes constant returns to scale and no other productive factors than capital and labour.

⁴ Griliches’ (1963) notion of economies of scale is discussed in part II of this paper.

countries; differences in ITC deflators result in enormous cumulative differences in measured factor accumulation.⁵ Thus, national differences in accounting methods contribute to differences in reported productivity residuals. Similarly, differences in tax systems can bias measured factors shares. In sum, the productivity residual, besides being the repository of measurement errors, also contains essentially arbitrary information resulting from different national fiscal and accounting conventions. Even among OECD member nations TFP comparisons are fragile; and data quality tends to be more problematic outside the OECD.

Jorgenson has continued research on TFP measurement; Jorgenson and Vu (2011), for example, calculate a variant of equation [6] that includes factor quality and quantity and distinguishes between IT (information technology) capital and other capital for 122 countries. Better growth accounting and factor measurement reduces the productivity residual: Jorgenson and Vu find that input growth (factor accumulation) is by far the most important determinant of economic growth: the share contributed by TFP varies from country to country and period to period but it generally accounts for between 15 and 20 percent of growth. This is almost the exact complement of Solow’s initial calculations that TFP accounts for over 85 percent of growth.

Although significant advances have been made in the conceptualization and measurement of TFP, our measurements, particularly at the aggregate level, lack precision; this limits their value in econometric analysis. Independent of the challenge of measuring TFP and quality changes in inputs and outputs is the challenge of explaining why change occurs. In the Solow growth model “technical change” is exogenous: it is caused by forces external to the economy. The problem for what is now called “exogenous” growth theory (the growth theory of the 1950s and 1960s) was to devise an aggregate-level economic theory of innovation, thus making “technical change” endogenous.

Knowledge, innovation and R&D spillovers

There was considerable economic research at the micro-level on the economics of innovation and inventive effort in the 1950s, but Arrow (1962) is regarded as the seminal contribution. Arrow identified the central problem in the economics of innovation: innovation is the generation of new information and information differs from most private goods in that it is indivisible, non-rival, and generally non-excludable. The table below reproduces the two-by-two typology of goods found in most economics textbooks. Inventions and innovations fall into the lower-right cell with public goods.

	Excludable	Non-excludable
Rival	Private goods	Common Goods
Non-Rival	Club Goods	Public goods

There is a cost of innovation to the innovator, but once a product is innovated, the cost of transmitting the new idea is approximately zero. And information, once transmitted, cannot be returned: it is difficult to “un-know” something once it is known, particularly if the information is useful or valuable. Because innovators cannot fully appropriate the gains from their efforts, Arrow argued that competitive market economies under-invest in innovative activity. This provides an argument for public support for basic (scientific) research. Patents, trademarks, licensing and trade secrecy are means and protecting creating property rights in information and preserving incentives for private innovation: there are institutional

⁵ See Schreyer (2002) on ITC bias and Schreyer (2001, 2009) on OECD measurement standards.

means of making innovations quasi-excludable. Also, even if information is free, it is not costless to interpret and put into use: reverse engineering and imitation require skills, effort and financing; some absorptive capacity is necessary. Private innovative activity can lead to aggregate productivity because adopting a new innovation costs less and is a less uncertain investment than the initial innovation. Innovation by one firm can result in positive externalities for other firms. Thus, Romer (1990: S72):

Once the cost of creating a new set of instructions has been incurred, the instructions can be used over and over again at no additional cost. Developing new and better instructions is equivalent to incurring a fixed cost. This property is taken to be the defining characteristic of technology.

Research on the economics of innovation and R&D expanded in the 1960s and 1970s. Griliches and his students at Harvard and collaborators through NBER were central to this research program.⁶ Griliches (1979) modified the Solow model, rewriting equation [3] as:

$$[7] \quad Y_{it} = A_t X_{it}^{(1-\beta)} R_{it}^{\beta} S_{it}^{\gamma}$$

In [7], for the *i*-th firm in an industry, the variable *X* represents all conventional inputs (capital and labour, properly measured and quality-adjusted); *R* represents the stock of R&D knowledge produced by the firm; and *S* represents the stock of ‘outside’ knowledge that the firm can draw upon, specifically, the R&D produced by other firms in the industry. Griliches imposed constant returns to scale on the firm’s own inputs (*X* and *R*) to emphasize that the knowledge spill-in of *S* can result in increasing returns. The summation of R&D in the industry, by construction, equals *S*; and the industry production function can be written as [8]:

$$[8] \quad Y_t = A_t X_t^{(1-\beta)} S_t^{\beta+\gamma}$$

This yields “an aggregate production function with the coefficient of aggregate knowledge capital being higher ($\beta + \gamma$) than at the micro level (β only), since at the aggregate level it reflects not only the private but also the social returns to research and development.”⁷ (Griliches (1979[1998:29]) Thus the R&D spillover was well-defined by 1979. The means of appropriating R&D had been investigated by Levin *et al.* (1987) and Levin and Reiss (1988) but a consistent incentive mechanism for undertaking R&D was still missing.

Endogenous growth theory

The key to endogenizing innovation in growth theory was the revival of Chamberlin’s concept of monopolistic competition by Spence (1976) and Dixit and Stiglitz (1977). Dixit and Stiglitz devised a model of consumer demand for variety: it provided an elegant mathematical representation of the proliferation of similar yet distinct branded products (e.g., shampoos, breakfast cereals, running shoes).

⁶ Much of this work is collected in Griliches (ed.) (1984) and (1998); Terleckyj (1974) and Mansfield *et al.* (1977) are important contributions from this period.

⁷ Griliches continues, setting out the research program that is now being realized: “The above formula provides a framework for reconciling micro and macro results in this area. Of course, this formula is rather simplistic and is based on a whole string of untenable assumptions, the major ones being: the assumption of constant returns to scale with respect to *X* and *R*, and the assumption of common factor prices for all firms within an industry. These assumptions could be relaxed. This would add a number of “mix” terms to the equation, indicating how aggregate productivity would shift if the share of, say, the larger firms, were to increase (as in the case of economies of scale).”

In their model, a firm developing a new variety can expect to recoup its costs and earn a reasonable profit from the premium charged to consumers. Such markets are not perfectly competitive; rather, each firm producing a distinct product enjoys a limited monopoly; thus, monopolistic competition.

New applications of their model were quickly realized.⁸ Krugman, Brander, Lancaster and Ethier (among others) employed the model to explain the existence of intra-industry trade, which had been an annoying anomaly in modern trade theory. Krugman (1979, 1980) showed the link between variety and scale economies.⁹ Ethier (1982) applied the demand for variety model to trade in intermediate inputs: when firms have a longer menu of inputs to choose from they have a greater chance of realizing productivity-improving input combinations; thus trade in intermediate goods is a potential source of TFP growth.

In the mid-1980s Romer (1986) and Lucas (1988) reignited interest in growth theory. Using competitive-market assumptions, both authors proposed a source of positive aggregate growth externalities. Lucas's model is based on externalities from increases in human capital; Romer's model is based on externalities from increases in the stock of knowledge. The conceptual synthesis came with Romer's (1990) article, "Endogenous Technical Change", which synthesized growth theory and monopolistic competition.

First, nonrival goods can be accumulated without bound on a per capita basis, whereas a piece of human capital such as the ability to add cannot. Each person has only a finite number of years that can be spent acquiring skills. When this person dies, the skills are lost, but any nonrival good this person produces ... lives on after the person is gone. Second, treating knowledge as a nonrival good makes it possible to talk sensibly about knowledge spillovers, that is, incomplete excludability. These two features of knowledge—unbounded growth and incomplete appropriability—are features that are generally recognized as being relevant for the theory of growth.¹⁰

Endogenous growth models explicitly include knowledge, a nonrival input, in the aggregate production function of the economy. This relaxes the constant-returns-to-scale assumptions of Solow (exogenous) growth theory: the aggregate economy exhibits increasing returns to scale.¹¹ Endogenous growth models have three sectors: the core is a monopolistic-competitive durable inputs sector that makes inputs that are used to make consumer goods by a competitive consumer goods sector (the second sector). The third sector is the competitive R&D sector; it is hired to design new technologies for the durable inputs sector. The monopolistic premiums earned by durable input producers provide the funds (and incentive) for contracting R&D work.

There are two variants of endogenous growth models – horizontal and vertical - based on the assumptions made about the innovation process. Horizontal models, such as Romer's, are based on Dixit-Stiglitz assumptions about product variety: innovation results in additional products that are used

⁸ Wilfred Ethier (2004: 149) describes the Dixit-Stiglitz model as providing a versatile "tool". Dasgupta and Stiglitz (1980a, 1980b) address the incentives for R&D under imperfect competition.

⁹ Economies of scale are a potential source of TFP growth but one excluded by the constant returns to scale assumption of Cobb-Douglas aggregate production functions used in growth theory.

¹⁰ Romer (1990: S75).

¹¹ Formally: let A be nonrival inputs and X be rival inputs, then $F(A, \lambda X) = \lambda F(A, X)$ and $F(\lambda A, \lambda X) > \lambda F(A, X)$; the production function $F(\cdot)$ is not homogeneous of degree one: it exhibits increasing returns to scale. Romer (1990:S76)

in combination with existing products.¹² An increasing number of products and an increasing stock of knowledge give rise to increasing growth externalities. Vertical models are based on a revival of Schumpeter's concept of creative destruction: innovations result in higher-quality products or processes that render earlier products and processes obsolete.¹³ In contrast to horizontal models where the key variable is the number or quantity of product varieties, the key variable in vertical models is the average quality of products and processes. Vertical and horizontal models are formally parsimonious: they represent the endogenous innovation in an aggregate model with two parameters: one for the degree of monopolistic competition and the other for stock of innovation, measured vertically or horizontally, respectively.

In both models, agents invest resources to acquire the exclusive ability to manufacture a new product. Moreover, the R&D activity generates inappropriable spillovers in both cases. In the variety-based growth model, the R&D externality is quite explicit. Each completed product development project lowers the cost of later R&D efforts. In the quality-based model, the externality is implicit. When one improvement project succeeds, other researchers can quit their efforts to achieve that same innovation and begin to work on the next improvement. In both instances we have assumed that by observing the results from one innovative success, researchers can learn scientific and engineering facts that are useful in their own research endeavours.¹⁴

The rate of growth of the stock of knowledge is the critical variable in the model. Here is Romer's (1990: S83) explanation of the innovation process driving the model.

If the researcher possesses an amount of human capital H_j and has access to a portion A_j of the stock of knowledge implicit in previous designs, the rate of production of new designs by researcher j will be $\delta H_j A_j$, where δ is a productivity parameter.

At the aggregate level, the rate of growth in A is the summation over all H_j engaged in R&D: $\delta H A$. That higher knowledge growth follows from more human capital employed in R&D is no surprise; what is novel is that a larger stock of knowledge, A , results in a higher rate of growth in A . Again, Romer (1990: S84):

Linearity in A is what makes unbounded growth possible, and in this sense, unbounded growth is more like an assumption than a result of the model. ... Whether opportunities in research are actually petering out, or will eventually do so, is an empirical question that this kind of theory cannot resolve. The specification here, in which unbounded growth at a constant rate is feasible, was chosen because there is no evidence from recent history to support the belief that opportunities for research are diminishing.

Subsequent empirical research has tested this proposition and found that there is no evidence for the strong scale effects hypothesized by the initial wave of endogenous growth models. A second generation of "semi-endogenous" growth models that includes diminishing return to R&D has

¹² Romer (1986, 1990) is the classic horizontal formulation. See also, Chapter 3 of Grossman and Helpman (1991a).

¹³ The initial formulations of the vertical model are Segerstrom, Anant and Dinopolous (1990), Aghion and Howitt (1992) and Chapter 4 of Grossman and Helpman (1991a).

¹⁴ Grossman and Helpman (1991b: 54).

developed in response; the research program continues but this goes beyond the topic of the current paper.¹⁵

Measuring R&D and R&D spillovers

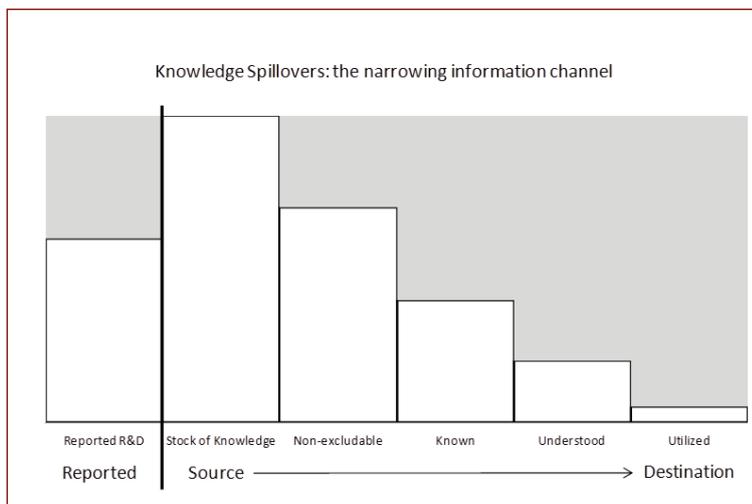
Equation [9], adapted from Levin and Reiss (1988), builds on Griliches' formulation [7] and helps to clarify the chain of empirical challenges in measuring R&D and R&D spillovers. The impact of R&D in a production function is determined by the firm's own R&D (R), the R&D of all other firms (S), and the degree of non-excludability of all others' R&D (β):

$$[9] \quad R^\alpha (\beta S)^\lambda \quad 0 \leq \beta \leq 1$$

The more other firms keep innovations secret or otherwise inhibit appropriation the lower the proportion of S that can be appropriated; this is represented by a lower value of β . This equation is a micro-level version of the Coe and Helpman (1995) equations discussed in the introduction. In the context of international R&D spillovers, the equation requires some elaboration. The coefficient β represents the proportion of S that could be appropriated by a given firm. The coefficient λ indicates how the stock of appropriable knowledge (βS) is utilized by the firm and realized as changes in cost reduction or TFP growth. In practice, it is difficult to measure the degree of non-excludability. Levin and Reiss (1988) had in-depth survey data with which to construct plausible micro-level measures of β , but such data are an exception. Also, there are other factors that reduce the effective stock of S to a given firm.¹⁶

$$[10] \quad R^\alpha (\beta_1 \beta_2 \beta_3 \beta_4 S)^\lambda \quad 0 \leq \beta_i \leq 1$$

These factors are enumerated below following the spillover channel from source to spill-in destination. These factors can be formalized in [10] as a series of information or transmission filters, their joint product is the effective information available to spill-in to a recipient firm. The figure below illustrates the process and it organizes the discussion.



Determining the relevant subset of knowledge

First consider S, the stock of knowledge: in principle the entire cumulative stock of human knowledge could be included in S; in practice one limits S to some relevant subset of all knowledge. Griliches (1998: 257ff) frames the problem as devising an appropriate weighting or distance function. For example, a firm producing paints is most likely to find the R&D of other

¹⁵ The seminal refutation is Jones (1995); for surveys of the subsequent debate see Jones (2005) and Gustafsson and Segerstrom (2010), which incorporates heterogeneous firms. On econometric issues see Durlauf *et al.* (2005).

¹⁶ Thus, in estimation, errors in measuring β influence the value of λ : the two parameters are difficult to distinguish.

paint producers of greatest potential value; varnish and lacquer producers would also qualify as close neighbors; and the segments of the chemical industry engaged in dyes and solvents would also be highly relevant. Other chemical segments and fluid-processing engineering are likely of slightly less value; but they would merit a positive weight. R&D in microprocessors and pharmaceuticals would probably be of little value and could be imputed a weight of zero. Thus, the relevant pool of outside R&D for a firm or industry is a weighted summation of neighboring industries. These weights can have some empirical grounding (input-output tables, econometric studies) but they inevitably contain arbitrary elements and value judgments.¹⁷

The reported stock of R&D (how this is measured is discussed in Appendix A) is merely an indicator of the stock of knowledge. Much innovation occurs beyond what is officially reported as R&D. Of the factors that limit the transmission of the stock of knowledge, the first, the degree of non-excludability has been discussed above (Levin and Reiss).

Contact and communication

The second limiting factor is that there must be some way for a firm to be aware that the outside R&D exists; this involves contact and communication. The earliest applied work on technological spillovers occurred in the 1940s in the U.S. Corn Belt where farm-level data was abundant and technological change was rapid and observable. Ryan and Gross (1943, 1950), rural sociologists, observed that farm operators differed in their willingness to try or adopt new farming practices. Few would adopt immediately; a minority would adopt in the second or third season; the majority would adopt later; and a small remnant would never adopt. This pattern was observed across communities and innovations and led to a general model of the diffusion of innovation (Rogers 1962). The social status of the early adopters and their success or failure influenced the rate of diffusion in the community (Rogers and Beal 1958).

Griliches (1957) provided an economic explanation of the diffusion process for the adoption of hybrid corn. Hybrid corn was a break-through innovation, resulting in large yield increases and improving the effectiveness of mechanized harvesting equipment. The new technology (genetics) was embodied in hybrid seed. Because data on the area planted to hybrids and traditional corn were available at the micro level Griliches could track the spatial diffusion of hybrid corn. He could observe hybrids developed in Iowa spilling over into Nebraska and into Illinois, for example.¹⁸

In empirical work communication is represented by some indicator of the probability or flow of contact. For example, geographic distance is often used; this is the basis of gravity models, it is also the basis of agglomeration theory, an idea that dates back to Alfred Marshall and which was revived in the 1990s.¹⁹ Measures of the value of bilateral or unilateral trade capture similar information; this is the variable favored in the literature following from Coe and Helpman (1995).

¹⁷ The work following Coe and Helpman's (1995) example uses only 'business sector' R&D; that is, they exclude public R&D. In agricultural economics, the convention is to include only agricultural R&D (public and private) although non-agricultural R&D obviously influences agricultural production: see, e.g., Alston (2002).

¹⁸ Traxler and Byerlee (2001) apply this method to wheat spillovers within India. For a survey of earlier industrial applications see Nasbeth and Ray (1974); also see Jovanovic and Rob (1989). Keller (2004) is a literature survey of international diffusion of technology. Conley and Udry (2010) focuses on information flows through social networks among pineapple growers in Ghana.

¹⁹ Krugman (1991) initiated the revival; Glaser et al (1992) provided an additional boost. Fujita and Thisse (2002) is the standard survey and text. On R&D spillovers and geography see Jaffe *et al.* (1993) and Audretsch and Feldman (2004). Greenaway and Kneller (2008) examine the relationships between exporting, agglomeration and productivity.

Human capital

There is enough evidence to give validity to the hypothesis that the ability to deal successfully with economic disequilibria is enhanced by education and that this ability is one of the major benefits of education accruing to people privately in a modernizing economy. - T.W. Schultz (1975: 843)

Human capital is the essential ingredient in innovation. Schultz, who invented the concept of human capital, viewed its primary economic function as the ability to adapt to change, to innovate and to learn. In Romer's model of endogenous growth the R&D sector has one factor of production: human capital. The capacity of a country to generate domestic R&D depends directly on its accumulation of human capital, in particular human capital above a critical technical threshold. Human capital is the key variable in innovation diffusion models: it is positively related to access to information, social status, and the capacity to comprehend and utilize information. Nelson and Phelps (1966) made explicit the relationship between human capital and technology adoption. Benhabib and Spiegel (1994, 2005) empirically verify the importance of these two roles of human capital in determining a country's TFP growth.

The fact that human capital influences growth through two channels, innovation and learning-adaptation, poses a conceptual problem for measuring R&D spillovers. The Griliches-Romer sense of an R&D spillover is limited to outside knowledge that is utilized in the formal R&D activities of firms and organizations. The goal of such R&D is the innovation of new products and inputs or improved production processes. Endogenous growth theory explicitly limits itself to formal R&D innovation; this is its source of TFP and its engine of growth. Learning beyond the walls of the R&D facility is embodied in the skill set of a particular person, which is rival-in-use (see Romer quotation above on page 8). Such learning is an augmentation of the stock of human capital and should be accounted as an increase in factor quality: it does not belong in the TFP residual. In practice, it is difficult to maintain this distinction; for example, Solow (1994: 177):

Bits of experience and conversation have suggested to me that it may be a mistake to think of R&D as the only ultimate source of growth in total factor productivity. I don't doubt that it is the largest ultimate source. But there seems to be a lot of productivity improvement that originates in people and processes that are not usually connected with R&D.

One interpretation of Solow's comment, consistent with the Griliches-Romer distinction, is that when non-R&D learning and adaptation results in new knowledge that is non-rival it may contribute to innovation and TFP growth; otherwise, when it is rival, it must be counted as an increase in human capital. The practical problem is that these distinctions are difficult to measure. In theory, an employee that spills-in outside knowledge and becomes more productive receives a proportionately higher wage (because wage equals marginal product in theory); the increased wage bill measures the increase in human capital and thus is not counted in the TFP residual. In practice, the increment in labour productivity is difficult for the employer to observe and individual wages are not so readily changed; very likely the increase in human capital is unmeasured or under-estimated and all or some of the increase is counted as TFP growth.

Although the role of human capital is universally recognized, there is less agreement over how best to measure and represent it at the aggregate level. The average years of schooling measure has intuitive appeal; but in practice there are national differences in reporting accuracy and school quality, which must be accounted for. Wößmann (2003) surveys the literature; Cohen and Soto (2007) advance an improved comparable measure of human capital, which includes differential mortality rates, an important factor that had been neglected.

Institutional context

The operating environment of the firm bounds its ability to utilize absorbed outside knowledge in production, procurement and distribution. Local and national customs, laws and regulations, for example, can inhibit or prohibit implementing innovations. The “new institutional economics,” advanced by Douglass North and Oliver Williamson in the 1970s and 1980s, has been incorporated into formal economic theory by younger economists. Daron Acemoglu is perhaps the one individual most responsible for this contemporary synthesis. Acemoglu argues that institutions, formal and informal, are key determinants of economic growth, primarily through their ability to encourage or inhibit the efficient allocation of factors. Barriers to matching factors of production with firms can trap resources in low-valued uses and limit the optimal division of labour. Thus the structure of legal systems, the efficiency of public administration and law enforcement, the level of inter-personal trust, and ethnic, racial, gender and age discrimination (formal or informal), among many other “non-economic” factors all influence the allocative efficiency of the economy.²⁰

Legal systems, for example, vary in the rights provided to minority shareholders and to creditors. Common law systems provide more rights than civil law systems and this allows greater opportunities for corporations to raise capital (La Porta *et al.* 1999). Bloom and van Reenen (2007) find that countries with a tradition of primogeniture have a high proportion of inefficiently managed family-owned firms. If the executive candidate pool is limited to eldest sons or immediate blood relations the likelihood of recruiting a competent executive is greatly diminished. Similarly, employment protection laws can result in labour market rigidities that reduce the ability of firms to adapt to changing market conditions and reduce the likelihood of innovation (Saint-Paul 2002, OECD 2002b, Botero *et al.* 2004.)

Conclusion Part I

The empirical literature on trade-related R&D spillovers since Coe and Helpman (1995) has updated and expanded the database, improved the econometrics,²¹ and tested alternative weighting systems for foreign R&D stocks. It has also incorporated additional explanatory variables, reviewed in the previous section. In estimation, the basic equation of Coe and Helpman (1995) (equation [2] above) is augmented with additional variables; the basic specification is:

$$[11] \quad \ln TFP_i = \beta_i^0 + \beta_i^d (\ln S_d) + \beta_i^f (m_i * \ln S_f^-) + \beta_i^h (\ln H_i) + \beta_i^Z (Z * \ln S_f^-) + \varepsilon_i$$

Stocks of R&D are represented by S where the subscript indicates d – domestic and f – foreign. The superscript on the foreign R&D term indicates the weighting system used to aggregate the R&D stocks of country I’s trading partners into a single R&D stock value. As in the 1995 paper, S_f is interacted with m_i , the ratio of imports to GDP. Human capital, represented by H, is invariably positive and significant and is now a standard in empirical application.

The additional variables are usually interacted with foreign R&D stocks, represented in equation [11] by the variable Z. The additional variables are indicators of institutional quality: the ease of

²⁰ Acemoglu *et al.* (2001) is the breakthrough article in this area. Acemoglu *et al.* (2005) is an excellent literature survey. Recent contributions directly relevant to R&D spillovers include: Acemoglu *et al.* (2006) and Acemoglu *et al.* (2007). Helpman (2008) is an edited volume devoted to institutions and trade. Barbosa and Faria (2011) is a recent overview of the links between institutions and innovation. Braguinsky *et al.* (2011) is a timely study of the impact of restrictive firing laws in Portugal on the size distribution of firms and stagnant productivity growth.

²¹ For more on innovations in panel data econometrics see Hsiao (2003) or Baltagi (2008).

doing business as measured by the World Bank index; the quality of tertiary education; the strength of intellectual property rights; and the origins of the legal system. These measures are usually divided into high and low scores, or into high, average and low scores, and included as dummy variables. The variables are generally significant and in the directions suggested in the literature. The ease-of-doing-business term, for example, interacts positively and significantly with human capital; Coe *et al.* (2009) interpret this result to indicate that a more business-friendly environment, *ceteris paribus*, increases the capacity of human capital to realize R&D spillovers.

There are several rival weighting schemes for foreign R&D stocks. A bilateral-import weighting [12] is the most common form used in the literature for aggregating foreign R&D stocks. A bilateral import-to-GDP measure [13] is advanced by Lichtenberg and van Pottelsberghe de la Potterie (1998) as an alternative. In the expressions below S_j is the stock of R&D of the J th country, M_{ij} are I 's imports from J , and Y is GDP.

$$[12] \text{ Bilateral imports: } S_i^{Bi-M} = \sum_{j \neq i} \frac{M_{ij}}{\sum_{i \neq j} M_{ij}} S_j$$

$$[13] \text{ Bilateral imports to GDP: } S_i^{m/Y} = \sum_{j \neq i} \frac{M_{ij}}{Y_j} S_j$$

Coe *et al.* (2009) find little difference between the import weighting-scheme regressions, but they favor import-weighting of external R&D stocks.²² Funk (2001) uses the form of equation [12] with bilateral export weights and finds that it performs as well as bilateral import weights. Xu and Wang (1997), argue that imports of capital goods may be more closely related to R&D spillovers than non-capital goods imports. They use the form of equation [12] but with capital goods imports in place of total imports and find that it yields a significantly larger coefficient than for the total imports ratio. Seck (2012), applying the analysis to developing countries, also finds that the capital-goods-import ratio performs better than total imports.²³ Seck (2012) also finds significant effects for equation [12] using FDI to total investment; and finds varying significance for private, public and university R&D stocks.

These weighting schemes are designed to represent the relative likelihood of spillover from the many potential R&D sources being aggregated. Contact is the most important component of this likelihood. All of the measures discussed above are partial measures of contact; thus one expects each of them to be significant if used as the sole indicator of contact. Importing and exporting each provide opportunities for contact; Grossman and Helpman (1991: 165ff) are explicit on this point in their discussion of “international information flows”.²⁴ They suggest that total (X+M) bilateral trade is the most appropriate measure and, following Arrow’s idea of learning-by-doing, that it should be calculated cumulatively. Grossman and Helpman, however, limit their discussion of information flows to activities linked directly to international trade. Commercial trade is only one of many potential means of contact between firms in different countries.

²² Coe *et al.* (1997) examines North-South R&D spillovers; which may account for preferring imports to exports.

²³ It is possible that the capital-goods weighting captures distortions in capital and credit markets in the importing country which are erroneously counted as TFP growth.

²⁴ The relevant passage is reproduced in appendix B of this paper.

The central argument of this part of the paper is that aggregate measures of national-level total factor productivity and national stocks of R&D capital are not precise measures; they are highly aggregated, averaged and smoothed in multiple dimensions. They are informative at an aggregate level, but one cannot expect them to support empirical analysis at a fine level of detail. The body of empirical aggregate-level work is consistent with the existence of international R&D spillovers and that they are positively related to many alternative measures of bilateral trade. That multiple indicators of international contact are correlated with R&D spillovers is consistent with there being multiple communication channels, but it is difficult to extract more information from aggregate data. Economists are overcoming the aggregate measurement impasse by shifting the analytical focus to the micro-level: this is the subject of part two.

PART II: R&D SPILLOVERS: MICRO-LEVEL ANALYSIS

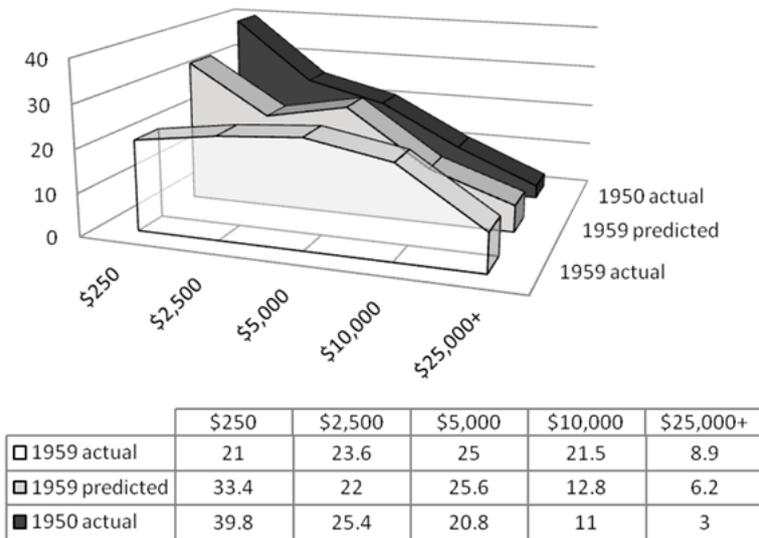
Introduction

In macro-level productivity analysis the implicit assumption is that the economy is one large firm or a set of identical representative firms all of which operate equally efficiently at the productivity frontier. When productivity increases, the efficiency frontier shifts and all firms move with the frontier. In reality, firms differ; few are at the efficiency frontier; the rest lag behind. In aggregate measures it is difficult if not impossible to distinguish between movements of the efficiency frontier and interior movements toward the frontier: both are measured as TFP growth. At the micro level it is possible to observe these differences.

The work of Zvi Griliches is central to the development of micro-level or heterogeneous-firm approaches to productivity analysis and R&D spillovers. In his diagnosis of the measurement problems common in aggregate-level productivity analysis Griliches (1963) included economies of scale at the firm level. Griliches observed that the U.S. agricultural sector exhibited rapid productivity growth in the 1950s; farms increased their use of capital services and purchased inputs and reduced the amount of labour employed; specifically, the number of farms and farm operators fell dramatically between 1950 and 1959. Agricultural economists had noted that most farms were too small, given the farms' capital and human capital endowments. Griliches' (1957) work on hybrid corn and Roger's (1962) work on diffusion of innovations documented the heterogeneity of apparently similar mid-western corn farms. Some farmers were simply better at farming than others; some were quick to adopt new technologies and plant varieties, others adopted much latter or never. Farm-operator heterogeneity is a source of observed productivity growth. Less-capable operators are more likely to leave full-time farming while more-capable operators are more likely to buy or rent-in the land from former farm operators. Thus there is a transfer of farmland and equipment from less-capable to more-capable farm operators.²⁵ The figure below is constructed from data in Griliches (1963: 339): it plots the distribution of commercial farms by sales class in the 1950 and 1959 U.S. Cencuses of Agriculture (these are the "actual" values; all figures are in 1954 dollars). It also plots a "1959 predicted" distribution: this is constructed by applying the 21 percent rate of U.S. agricultural productivity growth rate observed for 1950-59 uniformly across the 1950 distribution. Contrasting the 1959 actual and predicted indicates that observed productivity change was not evenly distributed: the share of larger farms grew more than predicted and the share of farms in the smallest class declined more than predicted.

²⁵ The reallocation of resources in U.S. agriculture from less-capable to more-capable managers continues; see Hoppe et al. (2010).

Griliches (1963): Distribution of U.S. Commercial Farms by Sales



When black box of aggregate productivity analysis is opened and one examines the dynamics of individual firms, one finds exit, entry and reallocation: less successful firms tend to contract or exit; more successful firms tend to expand; new firms emerge; and most factors in the industry are reallocated within the industry. Productivity growth is not scale-neutral; it is not uniformly distributed across incumbent firms. Measuring and understanding these processes is the core of heterogeneous firm analysis.

Micro-level analysis of productivity

Micro panel data

Heterogeneous-firm analysis developed to explain the commonly observed distributions of the size, productivity and growth of firms within industries. These consistent patterns attracted the attention of statisticians and economists as quality data became available in the late 1800s. Early research focused on devising plausible stochastic (random) processes that would generate the observed size distributions and rates of firm exit, entry and growth.²⁶

It required the construction of panel or longitudinal data sets, where a specific cross-section of individual firms or respondents is surveyed in multiple periods, to observe and analyze individual firm dynamics, particularly the decisions to expand or contract production and to enter or exit an industry. The pioneering work on panel data econometrics involved panels of farms: Hoch (1955, 1958, and 1962) worked with a Minnesota farm panel; Mundlak (1961) worked with an Israeli farm panel. They proposed what is now called the fixed-effects model; this is a way of controlling for unobserved individual differences in farms, firms or individuals. Hoch referred to the unobservable variable as differential managerial ability or farm-specific technical efficiency; Mundlak likewise referred to individual differences in management and viewed fixed effects as a means of estimating production

²⁶ See Sutton (1997) for a literature survey; this was an active issue in economics in the 1950s and early 1960s: Adelman (1958), Simon and Bonini (1958), Mansfield (1962).

functions “free of management bias.”²⁷ What Hoch and Mundlak identified and attempted to control for is the heterogeneity of firm (or farm) productivity. It was obvious by 1960 that the only way to understand the fundamental, firm-level basis of productivity change was to construct and analyze panel data sets.

Constructing panel data sets is expensive and often can only be accomplished by public authorities. One of the challenges to applied research is the confidentiality of much official survey and census data. In the United States the NBER has cooperated (after much negotiation) with the U.S. Bureau of the Census which allows analysis to be published so long as the confidentiality of individual respondents is preserved. National governments differ in the degree and terms of access offered to researchers.²⁸ Data development, data access and panel econometrics developed in conjunction in the 1970s and 1980s.

By the early 1990s several stylized facts had been established.²⁹ First, there are large and persistent differences in firm productivity within industries. Using Markov transition matrices, the common pattern observed is that highly productive firms tend to remain highly productive and less productive firms tend to remain less productive. Second, the risk of exit is inversely related to the level of productivity: that is, less productive firms are the most likely to exit. Third, entering firms, on average, have productivity levels similar to the average of incumbent firms. Fourth, despite considerable exit and entry of firms, factors (particularly labour) tend to remain in the industry; that is, former employees of exiting firms generally find employment with surviving or entering firms in the same industry: the reallocation of factors from less to more productive firms is a primary driver of productivity change at the industrial level. There are, of course, inter-industry factor movements, particularly out of industries in secular decline and into industries in secular growth; but these flows are small relative to intra-industry reallocation.

A series of papers by Bernard and Jensen (1995, 1999) analyzed micro data on firms that export, focusing on firm productivity and entry-exit dynamics. Several important stylized facts emerged. Exporting (and importing) is relatively rare: few firms engage in it.³⁰ Firms that trade tend to have higher productivity than firms in the same industry that only produce for the domestic market. Firms that trade also tend to be larger and better capitalized than non-exporters. These empirical findings, among others, are forcing fundamental changes in international trade theory. Prior to this seemingly anomalous empirical evidence, international trade theory implicitly assumed that all firms in an industry were homogeneous; if Portugal has a comparative advantage in wine production, relative to English cloth, then all wine producers in Portugal were implicitly assumed to export to England. In fact, the distribution of exporting is concentrated in a minority of relatively high-productivity firms. There is a now a “new new” trade theory, heterogeneous-firm trade theory, that is attempting to construct a

²⁷ The bias resulting from the endogenous choice of inputs was first identified by Marschak and Andrews (1944). For more on the development of panel data econometrics see Nerlove (2002).

²⁸ Norway is unusually open in this regard and early empirical work used Norwegian data: Griliches and Ringstad (1971).

²⁹ The seminal article is Baily et al. (1992). Bartelsman and Dom (2000) is an excellent survey on panel data analysis particularly as it relates to total factor productivity. Syverson (2011) covers the many subsequent innovations in the literature. Tybout (2000) surveys the literature on developing countries.

³⁰ Bernard *et al.* (2009): in 2000, 3.1 percent of U.S. firms exported; 2.2 percent imported; and 1.1 percent both imported and exported. Trade by value is highly concentrated: the largest 1 percent of exporters accounts for 81 percent of export sales; the analogous share for importers is 78 percent.

theoretical framework that corresponds to the empirical patterns about firms that trade.³¹

One of the key questions in this emergent field, and the focus of this paper, is what accounts for the positive correlation between firm productivity and export status. Does exporting (or importing) cause a firm to become more productive? Perhaps trading firms have more contact with other countries and this provides greater access to information and thus leads to greater R&D spillovers which are then manifest in higher productivity. Or, perhaps trading firms are able to exercise greater market power than non-traders or to realize economies of scale not available to non-traders. These are the leading hypotheses underlying the causal arrow from trade to productivity. The causal arrow running the opposite direction, from productivity to trade, rests on selection, or self-selection: high-productivity firms are more likely to become exporters than low-productivity firms. All of these hypotheses are plausible and they are not logically mutually exclusive: they could be valid simultaneously, and this complicates empirical hypothesis testing. In the empirical literature reviewed below the trade-causes-productivity path is modeled as a series of learning effects and the productivity-causes-trade path is modeled as a series of selection effects.

The next section discusses some of the empirical tools employed in the analysis of firm-level dynamics within an industry; the extensions of these methods to firms that trade are discussed in the subsequent sections.

Heterogeneous firm dynamics: empirical methods

There are two stages in the analysis of productivity with panel data. First, one needs to measure productivity for each firm in each time period. Given these measures one can then analyze firm productivity dynamics; this section discusses productivity dynamics first and then turns to the measurement of productivity at the firm level.

Productivity decomposition, learning and selection

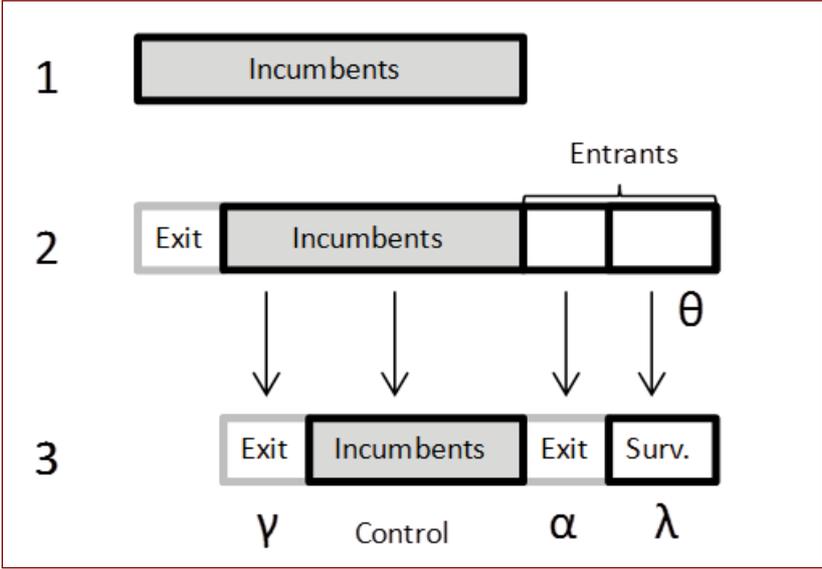
Productivity decomposition identifies and measures the relative contribution of different sources of productivity change in an industry. There are several productivity decomposition methods³² but their common starting point is the fact that industry-wide productivity is the weighted average of the productivities of the individual firms in the industry. Industry-level productivity change can be decomposed into five components. Three components relate to surviving, continuing firms: 1) changes in within-firm productivity (holding market share constant); 2) changes in market shares (holding productivity constant); and 3) the interaction of productivity and market share changes. These effects are often called within-firm, between-firm, and cross-firm effects, respectively. The fourth and fifth components are the contributions of entering and exiting firms, respectively. Both are measured relative to industry-level productivity in the base or end period because there is insufficient data to calculate within-firm productivity change.

Regression analysis can be employed to distinguish selection and learning effects. The dependent variable is firm productivity (P_{it}) for firm (i) at time (t). The panel of firms is unbalanced, that is, there are exits and entries. The focus is on all firms that were active at the time of survey 3. Entrants, in this context, are those firms that entered the industry after survey 1 but before the survey 2; and exits are

³¹ The last decade has been turbulent for international trade theory: Redding (2010) is a recent, concise survey of theoretical developments. Bernard *et al.* (2009) provides a statistical 'portrait' of U.S. trading firms; it gives due attention to importers. The literature to date has emphasized exporters.

³² Melitz and Polanec (2009) is a recent critical survey. Baily *et al.* (1992) is seminal work in this area and provides the basis of the discussion in this paragraph.

firms that left the industry after survey 2 but before survey 3. The figure below illustrates the different subsets of firms distinguished by dummy variables in the regression equation. The dummy variable for year controls for average productivity changes between surveys 2 and 3.



$$P_{it} = \beta + \gamma IncmbExit_{it} + \alpha EntExit_{it} + \lambda EntSurv(3)_{it} + \theta EntSuvr(2)_{it} + \delta Year_{it} + \epsilon_{it}$$

This example is based on the analysis in Foster *et al.* (2006: 754ff.), which finds the following, where strict inequalities indicate a significant difference (F-test):

$$\alpha < \gamma < 0 \leq \theta < \lambda$$

These findings are consistent with the emergent stylized facts about heterogeneous firm dynamics. The non-strict inequality ($0 \leq \theta$) indicates that the average productivity of entrants that survived to period 3 is greater than the average productivity of incumbents (in period 2) but the difference is not statistically significant. The inequalities ($\alpha < \gamma < 0$) indicate: 1) that entrants that exit had a significantly lower average productivity (in period 2) than incumbents who exit; and 2) that incumbents who exit had a significantly lower average productivity (in period 2) than surviving incumbents (surviving incumbents are not distinguished by a dummy variable and serve as the control for the regression). These results are evidence of selection: the average productivity of the industry increases because lower-productivity exiting firms are replaced by more-productive expanding incumbents and new entrants. The inequality ($\theta < \lambda$) indicates that surviving entrants had a significantly higher average productivity in period 3 than they did in period 2. This last result is evidence of learning, even after controlling for year effects. Foster *et al.* (2006: 755) remark: “This pattern implies the surviving entering cohort exhibits more rapid productivity growth than more mature surviving incumbents over this same period. That is, these results are consistent with post-entry learning-by-doing effects playing a nontrivial and statistically significant role.”

Measuring firm productivity with panel data and selection

Estimating productivity at the firm level with panel data involves many of the same measurement problems encountered at the aggregate level; it also introduces problems that are masked by aggregation:

simultaneity bias, selection bias and omitted price bias.³³ The binding constraint is always the breadth and quality of the data. Firm-level data is often insufficient to estimate or calculate TFP. Data are often collected on revenue or sales; without sales price data revenue cannot be converted into physical unit terms. Multi-product firms require a firm-specific price index. Similar measurement problems exist in factor measurement. A common solution is to use labour productivity.³⁴

Simultaneity bias emerges because a firm’s choice of variable input use is influenced by factors not observable to the econometrician, such as the firm’s knowledge of its productivity level and its expectations of market conditions. In estimating firm-level production functions the error term can be correlated with variable inputs and thus bias the estimated coefficients; this, in turn, results in biased productivity measurements.

Selection bias exists because the set of firms one observes is an outcome of a selection process. One does not observe firms that have chosen to exit the industry; nor does one observe firms that have chosen not to enter the industry, this is an important consideration when the focus of analysis turns to firms that trade as they may differ from firms that produce but do not trade.

Olley and Pakes (1996) devised an estimation algorithm that addresses both the simultaneity and selection biases. The intuition behind the Olley-Pakes algorithm is that one can use observed information about a firm’s current and past investment decisions (net changes in capital) as an indicator of the firm’s unobserved productivity level. Including this derived measure in the regression takes care of the simultaneity bias. And, based on the plausible assumption that the probability of survival is increasing in productivity, the derived indicator also accounts for selection bias.

Underlying Olley-Pakes and its various extensions is a dynamic optimization process governing a firm’s investment programme and its discrete choices about entry and exit. This provides a transition to a discussion of the discrete choice of whether a non-exporting firm becomes an exporter and vice-versa. The starting assumption is that there is a fixed sunk cost (F) for a non-exporter to become an exporter. The table below collapses and simplifies expected net present value calculations into a set of inequalities. The two rows contrast the decisions of exporters and non-exporters, both of which are assumed to be incumbent producing firms in the same industry. In the exit column, both exporters and non-exporters choose to exit if expected profits (π) are negative. If expected profits are positive exporters continue as exporters. Non-exporters with positive profits have the option of becoming exporters, but this is only economically rational if the expected discounted flow of future profits exceeds the fixed cost of becoming an exporter.

Initial state	Exit	Remain	Export
Exporters	$\pi < 0$	$0 < \pi$	
Non-Exporters	$\pi < 0$	$0 < \pi < F$	$F < \pi$

³³ Van Beveren (2012) is a recent survey of the econometrics of TFP estimation with panel data; it is a good starting point for new variants of the Olley-Pakes algorithm.

³⁴ Foster *et al.* (2008) explore three variant measures of firm-level TFP and evaluate their strengths and biases.

In reality these decisions are not as crisp and mechanical as portrayed in the table or in the models on which it is based (e.g., Roberts and Tybout (1997)); these decisions are often modeled empirically in a probit framework. In this context, a probit model would express the probability that a non-exporter becomes an exporter as an increasing function of value of $[\pi-F]$, among other variables.

Heterogeneous trading firm dynamics

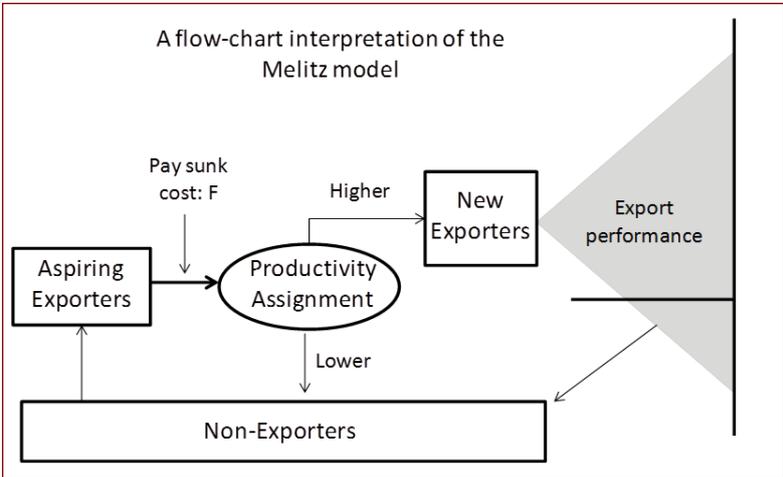
The theoretical watershed in heterogeneous-firm trade theory is the Melitz (2003) model. The Melitz model is driven by a selection process similar to those described above (Foster, *et al.*, Olley and Pakes, and Roberts and Tybout) and embedded in a differentiated-product trade model. The Melitz model is the starting point for contemporary trade theory; this is an active area and there are numerous extensions. This section provides an intuitive description of its selection process with an emphasis on the model’s implications for the analysis of learning and spillovers.

The figure below is a flow chart of the transition to exporter status. The central feature of the model is represented by the oval labeled “productivity assignment.” Aspiring exporters select themselves from active non-exporters in a given industry. An aspiring exporter must incur a fixed sunk cost to be assigned a productivity level. Each aspirant’s productivity assignment is randomly drawn from a probability distribution. Aspiring exporters are those who have made the expected net present value profit calculation and found that its expected value is positive.

When productivity assignments are made, aspirant exporters re-calculate their expected profits: those assigned lower productivities withdraw from the process and remain as non-exporters (less the fixed cost incurred). Those assigned higher productivities proceed to become new exporters and begin exporting. Their profitability in the first exporting period is determined by their productivity assignment and a market-wide stochastic element. The shaded cone at the right of the figure is meant to represent a distribution of profit outcomes for the cohort of new exporters. Each firm decides, based on its realized performance, whether to continue exporting. The horizontal line illustrates a potential cut-off point, a threshold below which firms revert to non-exporter status.

The exit threshold is not fixed: it is determined by market conditions. For example, an export “boom” induces a rapid increase in output as well as proportionally rapid increases in derived factor

demands. This drives up factor rental rates. Depending on the relevant elasticities in product and factor markets a boom could raise or lower the exit threshold. The important point is that given parameter values one can simulate the evolution of the productivity distribution of the export industry. Similarly, one can derive the expected changes in the industry distribution following innovations in trade policy (e.g., tariff changes, domestic and or foreign) and



innovations in factor and product markets.³⁵ Any argument in an exporting firm's profit function can shift the exit threshold.

The original (2003) Melitz model assumes that aspiring entrants are identical: the productivity assignment mechanism (i.e., the underlying Pareto distribution) is the initial source of exporter productivity heterogeneity; this initial distribution is then truncated (by low-productivity immediate exits) and modified by subsequent rounds of market selection and entry. These assumptions simplify the model: in reality firms self-select into becoming exporters.³⁶ The figure above is constructed to include self-selection: this is represented by the arrow linking non-exporters and aspiring exporters.

Selection, learning, innovation and exporting firms

Clerides *et al.* (1998) helped frame the empirical question of the relative importance of selection and learning in accounting for the observed positive correlation between a firm's export status and its productivity. From panel data they constructed characteristic cost trajectories for non-exporters and entering, exiting and continuing exporters. They found that the unit costs of entering exporters decline for several periods *before* they start exporting. At entry, entrants have approximately the same costs as continuing exporters and they are more efficient (lower-cost) than exiting exporters. This is the same pattern found by Foster *et al.* (2006) discussed above. They found little significant evidence that the post-entry cost profiles of entrants and continuing exporters differ, that is, little support for post-entry learning.³⁷ If there is no post-entry learning then the narrative can be reduced to a pure selection-driven stochastic process, such as the Melitz model: if a non-exporting firm experiences a random productivity shock that shifts its productivity above a critical threshold then the firm starts to export with probability 'p'; otherwise, it does not export. One can estimate the parameters governing the distribution of productivity shocks and closely replicate the observed distributions of exporting firms by productivity and by size.

López (2005, 2009) advances an alternate reading of the cost trajectories of entering exporters. He argues that the decision to export occurs several periods before a firm actually exports. Firms planning to export take productivity-improving actions prior to exporting. López refers to such actions as learning-to-export, as distinct from learning-by-exporting, which refers to post-entry increases in firm productivity.³⁸ The López narrative asserts two causal channels between exporting and productivity: pre-entry and post-entry learning. Self-selection still plays an important role in this story because some firms choose to plan to export and others choose not to.³⁹

³⁵ For example, Trebler (2004), examining the effects of Canadian tariff reductions, finds that high-productivity exporters expand market share and low-productivity firms contract or exit. Bernard *et al.* (2006) find a similar distribution of selection effects following a decline in transport costs.

³⁶ The simplification is effectively parsimonious as the observed distribution of exporting firms corresponds closely to a Pareto distribution. The Pareto distribution is governed by one parameter; this allows changes in the distribution of firms to be tracked in one dimension. This is a powerful simplification and it seems to be emerging as "useful tool" in theory construction similar to the Dixit-Stiglitz mechanism.

³⁷ Bernard and Jensen (1995) is the seminal paper in this area. Bernard and Jensen (1999) find very strong selection effects and no learning effects for U.S. exporters.

³⁸ Learning-by-exporting is analogous to Arrow's learning-by-doing. Following Arrow, the proper measure of a firm's export experience for learning-by-exporting is not the length of time since entry (as it is usually measured in the empirical literature) but the cumulative volume of the firm's exporting activity at a given moment. Fernandes and Isgut (2005) on Colombia is an exception: they follow Arrow and use a cumulative measure.

³⁹ A major empirical problem is that one must observe when a firm decides to start preparing to export. To make valid comparisons one must also observe those firms that make preparations to export but never export. On self-selection in exporting see the Olley-Pakes discussion on pages 21-22 above.

The standard empirical approach to investigating learning and selection is analogous to the analysis outlined above by Foster *et al.* (2006). An important difference is that the primary movement is between non-exporting and exporting. One can generally assume that exporting firms can be observed as active non-exporting firms before they become exporters and, if they cease exporting, that they often remain active in the domestic market rather than cease operations completely. Consequently it is possible to measure an exporting firm's productivity pre-entry, post-entry (if it survives) and post-exit (if it does not survive as an exporter). Such firms can be compared with cohorts of non-exporters who do not become exporters. A sizeable empirical literature exists on this topic and there are two good literature surveys: Wallace (2007) and Greenaway and Kneller (2007); and in 2008 ISGEP (International Study Group on Exports and Productivity) published a pooled study of 14 countries. There is clear evidence of an "exporter productivity premium": exporters have higher productivity than non-exporters; there is also uniformly strong evidence of self-selection, consistent with pre-entry learning; the evidence for post-entry learning is weak, however.

One limitation of the standard approach is that it does not fully control for selection effects; it measures the difference between the average outcomes for non-exporters and new exporters. The ideal counterfactual for an exporting firm is the performance of the identical firm had it not become an exporter, an alternative reality that one cannot observe. A close approximation is to match otherwise similar exporters and non-exporters and examine the pair-wise differences in productivity trajectories. Arnold and Hussinger (2005) use matching and find zero post-entry learning effects for German exporters. Girma *et al.* (2004) use matching and find positive evidence of post-entry learning for U.K. exporters. Yang and Mallick (2010) use matching and find significant post-entry learning effects in the second year of exporting for Chinese exporters. Generalization are not possible without a large body of similar studies, the limiting factor to the expansion of this promising empirical research program is the availability of quality micro data.

The statistical portrait of U.S. trading firms by Bernard *et al.* (2009) merits intensive study: the facts are illuminating and shed light on learning-to-export. The table below shows the distribution of exporting firms in 2000 by number of products exported and by the number of export destinations.⁴⁰ The typical or modal U.S. exporting firm exported one product to one country. The median is two products and one destination country. U.S. exporters are not atypical; similar distributions are found for France and for Colombia.⁴¹ In Colombia, where there is data on the number of customers an exporting firm has, the typical exporter exports one product to one customer in one country.

Number	Number of Products Exported		Number of Export Destinations	
	Firms%	Value %	Firms %	Value %
1	38	1	57	4
2	16	1	15	2
3-4	16	2	12	3
5-9	15	4	9	5
10	16	93	8	86

⁴⁰ Bernard *et al.* (2009) tables 14.4 and 14.6; Data are for 2000 and are rounded. The distributions for importing firms are similar to those for exporting firms.

⁴¹ See Eaton *et al.* (2011) and (2007), respectively. Relatively simple stochastic processes can simulate these distributions: Chaney (2011) is an indication of the direction of this line of research.

These statistics indicate that most entering exporters probably have a relationship with a specific foreign customer or customers prior to exporting. Consistent with the learning-to-export argument, such firms are likely to have made investments or undertaken some process or managerial improvements to meet contractual specifications for the new client. Some firms may take a speculative approach and attempt to export without a prior sales commitment; but this is rather risky given the costs involved.

Recent studies examine the distinction between exporting generally and exporting to a specific destination or destinations. Trofimenko (2010) and Park *et al.* (2010) find that exporting-firm productivity gains are more likely and stronger for firms exporting to rich, developed countries than for firms that export to lower-income developing countries. These findings are consistent with learning-by-exporting, technology-transfer, and R&D spillover narratives; however, they are also consistent with self-selection: if exporting to richer markets requires a higher fixed entry cost then the observed differences by destination can be largely a result of selection.

The mixed evidence of post-entry gains in productivity for exporting firms is consistent with international R&D spillovers; but it does not provide direct or conclusive evidence. There is a small literature that examines the relationship between exporting status and R&D at the firm level. Several studies find evidence that firms invest in innovation (R&D) prior to exporting and in anticipation of trade liberalization.⁴² But the dominant direction of causality is from innovation to exporting: engaging in the former increases the probability of subsequently engaging in the latter. Once again, self-selection is at work here. Damijan *et al.* (2008) is exceptional in that it finds evidence of export-induced innovation. The study uses nearest-neighbor matching to control for selection effects; but the far stronger influence in this study, selection, is from innovating to exporting.

One reason why little evidence of a link between exporting and innovation (R&D) is observed is because few firms engage in either activity. Typically it is larger, well-capitalized, and relatively efficient firms that self-select into exporting and engaging in R&D; very few firms engage in both activities.⁴³ The population of new exporters, in contrast, is largely comprised of small and medium sized firms that are unlikely to engage in R&D. The methods reviewed above treat individual cases equally; dummy variables capture the difference between the simple (un-weighted) average productivity changes of new exporters and non-exporters, for example. Thus, because of the low frequency of firms that export and innovate, one would expect the observed median R&D effect to be zero and the observed average effect to be insignificantly different from zero.

Selection, learning and innovation at the industrial and national level

The Melitz model concerns the entire distribution of firms, not merely those that export. The previous sections focus on selection into exporting; but there is also a lower productivity threshold that determines entry to and exit from the domestic market. Both thresholds are endogenous to market conditions. Trade liberalization, for example, by increasing import competition and raising potential returns to exporting and importing, shifts the thresholds for trading but also shifts the exit threshold for non-exporters.

Competition and the distribution of x-efficiency

Leibenstein (1966) introduced the concept of x-efficiency to describe the common empirical observation that firms do not operate as efficiently as assumed in economic theory. Heterogeneous firm theory has

⁴² Aw *et al.* (2008, 2011), Bustos (2011), Lileeva and Trefler (2010). The argument is that there is a simultaneous self-selection into (additional) innovation and exporting.

⁴³ On the distribution of R&D among firms see, for example, Cohen and Klepper (1992).

revived and reinforced Leibenstein's insight. Bloom *et al.* (2010a) measure the managerial practices employed by firms in 16 countries, finding a wide range of management-quality scores within each country, similar to the dispersion of firm-productivity scores found in other studies; and significant differences between mean country scores. They explain much of the variation in management quality by the degree of competition a firm faces, whether it is engaged in trade (trading firms have higher quality) and the formal education of managers. The degree of decentralization (how much discretion lower management is allowed) is also an important determinant; decentralized management is highly correlated with the degree of interpersonal trust in society; decentralized management is greater in common law countries than in civil law countries, for example.

Syverson (2004) and Bloom *et al.* (2010b) examine the relationship between competition and firm efficiency under conditions of spatial market power in two non-tradable markets: in cement production and public hospitals, respectively. They find that firm efficiency is largely determined by the proximity of competing firms. Bloom *et al.* (2011a) take a clinical trial approach to management quality in textile plants in India. Randomly selected firms were provided management consultant services for free for one month; control firms did not receive these services. The consultants' recommendations primarily concerned three issues: quality control, inventory management, and the physical flow of work. The recommendations were not new ideas, they could be found in any management textbook published after 1960. Not all treated firms acted on the recommendations, but the average treatment effect was an 11 percent increase in productivity.

Hsieh and Klenow (2009) examine the distribution of x-efficiency (as measured by the misallocation of capital and labour) at the firm level in India, China, and the United States. They calculate what the net productivity gain would be if the efficiency distribution of firms at each industry at the four-digit level in China and India were to shift to the associated U.S. distribution. They note that this is essentially a Melitz model without trade liberalization: rather than measuring the impact of trade liberalization on the distribution of firms (as in the articles reviewed below) they estimate the maximum potential change. The central estimate is a 40-percent and 50-percent increase in aggregate TFP, for China and India respectively.

All of these studies above confirm and refine the earlier work of Nickell (1996) on competition inducing higher firm-level productivity. There is emerging line of research that takes trade liberalization events as natural experiments that increase competitive pressures in an economy and contribute to TFP growth. This literature is the focus of the next section.

Liberalization, competition and innovation

There is a growing empirical literature that treats trade liberalization events as natural experiments. The common causal path is that trade liberalization increases competitive pressures in the domestic economy and induces domestic TFP growth, primarily through selection effects.

Pavcnik (2002) examines trade liberalization in Chile using firm-level data. She finds that surviving firms in import-competing sectors realized significantly higher productivity gains (4.6 percent annually) than in export-oriented (3.6 percent annually) and non-tradable (0.1 percent annually) sectors. Exit is an important factor: in aggregate, 70 percent of productivity gains can be attributed to the reallocation of factors among firms. She also finds that the relative lack of barriers to exit in Chile facilitated factor reallocation; restrictions on bankruptcy, plant closing and redundancy inhibit productivity gains.

Trefler (2004) examines the impact of the Canada-U.S. FTA on Canadian-firm labour productivity. Import-competing industries experienced increased firm exit and reduced employment. This was more

than offset by unusually high rates of labour-productivity growth among surviving firms. Trefler finds that about half the productivity gain can be attributed to exit and inter-firm factor reallocation and half to within-firm increases in technical efficiency.

Amiti and Konings (2007) examine the impact of tariff liberalization on Indonesian firms. They construct firm- and industry-specific measures of the change in the effective rate of protection effected by liberalization: that is, specifying weighted indicators of both output and input tariff changes. They find that reductions in input tariffs generate substantially higher within-firm productivity gains than reductions in output tariffs. They also control for industry-level competition by including a Herfindahl index alone and interacted with the output-tariff change variable. They find that within-firm productivity growth is negatively related to the degree of industry concentration; output tariff reduction induces productivity growth only for firms in competitive industries. This finding does not conform to the assumption of the endogenous growth theory that imperfect competition is positively associated with innovative activity and within-firm productivity gains.

Lileeva and Trefler (2010) examine how the Canada-U.S. FTA (Free Trade Agreement) influenced the export-entry decision by Canadian firms. They find that export entrants are disproportionately lower- and mid-productivity firms that were induced by import competition to invest in new technology, make process innovations and improve management; such firms undertake these changes at significantly higher rates than similar non-exporting firms. Export entrants not only exhibit within-firm productivity gains but also expand output volume and sales to the domestic (Canadian) market. This expansion, in turn, increases domestic competition and raises the exit rate.

Eslava *et al.* (2009) examine the impact of tariff liberalization in Colombia. Their Colombian data permits estimation of firm-level TFP. Their results are consistent with the other studies reviewed in this section. What is novel is that they use their statistical results to simulate the counterfactual rate of exit (as a function of firm TFP) using pre-reform tariffs. Comparing the counterfactual with realized exit provides a measure of the change in the exit threshold: liberalization causes a significant increase in the minimum threshold of firm TFP required for firm survival.

Bloom *et al.* (2011b) examine the impact of China's entry into the WTO on firm-level innovation, exit, and productivity in the EU-12. Consistent with the other studies, greater exposure to Chinese imports induces greater levels of innovation, investment in information technology and improvements in managerial practices. Chinese imports also led to differential selection: the incidence of exit increased for lower-technology firms relative to higher-technology firms, and the latter increased domestic market share. One important aspect of this study is that it underscores the role of increased competition in inducing innovation and productivity growth. Because the trade flow is South-North, from China to the EU, one expects there to be very little of the R&D spillovers or technological transfer assumed to exist when the import shock is largely North-South.

Similarly, Iacovone *et al.* (2011) examine the impact of Chinese imports on Mexican firms. They find a similar differential response: import competition induces a greater productivity response from ex ante higher-productivity firms than from lower-productivity firms. Moreover, the increased productivity does not result from innovation or R&D spillover, but from investing in improvements in quality control and inventory and personnel management. These are the same mundane sources of intra-firm productivity growth adopted by Indian textiles firms in Bloom *et al.* (2011a).

Bloom *et al.* (2011b), Iacovone *et al.* (2011) and Amiti and Khandelwal (2009) utilize liberalization events to investigate empirically the inverted-U relationship between competition and innovation

advanced by Aghion *et al.* (2005) and Acemogul *et al.* (2006). The inverted-U is the vertical summation of two opposing effects.⁴⁴ For firms far from the technology frontier, an increase in competition reduces the incentive to innovate because the gains from innovation are likely to be reduced by entry, but for firms near the technology frontier an increase in competition encourages innovation.

Heterogeneous trading firms and related-party trade

About two-thirds of the value of world trade consists of intermediate products. About one-third of world trade is intra-firm trade, most of which is of intermediate products. Heterogeneous-firm trade theory has drawn on organizational theory to explain the increasing international fragmentation of production and the growing importance of related-party trade. FDI is a potential channel for R&D spillovers; there is an empirical literature on this topic, reviewed by Görg and Greenaway (2004), but with inconclusive findings. This short section merely outlines the theoretical framework of this emerging strand of research.⁴⁵

The Coasian theory of the firm views the firm as a nexus of contracts and the key question is whether to make or to buy inputs. If inputs are readily available in liquid markets, they can be purchased as needed with only market-price or delivery risk. If markets are not sufficiently liquid or the inputs are in some way specialized then a contractual agreement is needed to assure supply. Contracts can be made at arm's length with another independent firm. But if contracting is not feasible, for example, if proprietary information cannot be revealed outside the firm, then production is done within the firm. International trade, when these relationships cross international borders, adds a second dimension: whether to contract at home or abroad. The international analogs are: anonymous non-related-party trade, foreign outsourcing, and vertical integration (foreign direct investment: FDI).

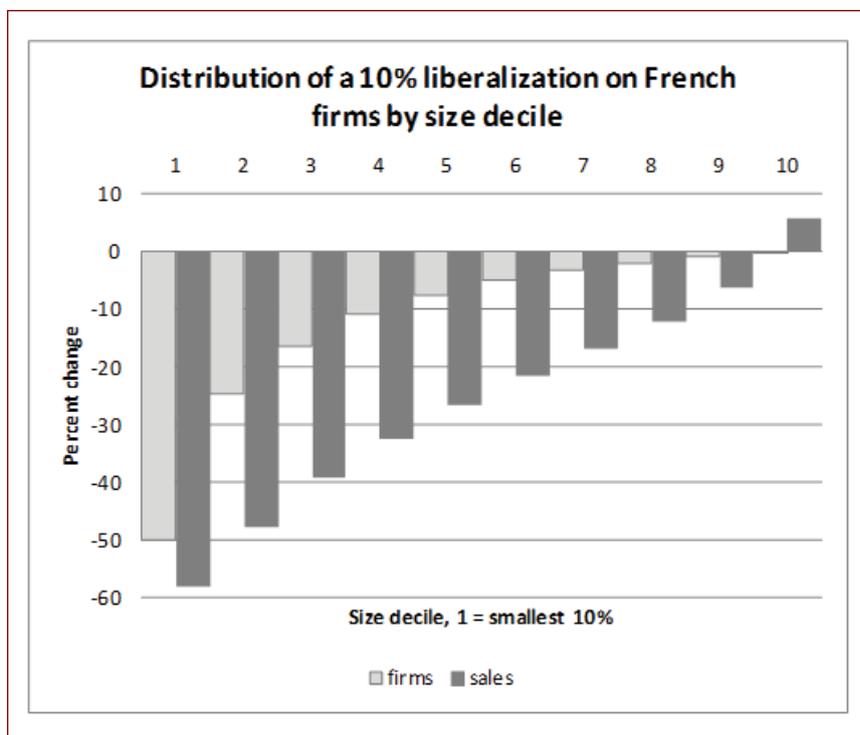
The 2-by-2 matrix below shows the union of the make-buy and home-abroad dichotomies. Firms face four alternatives for contracting specialized inputs. As in the Melitz model, each mode requires a fixed cost; there is a hierarchy of fixed costs and firms select the optimal mode based on their relative productivity. The least productive firms only outsource domestically; the next tier is domestic vertical integration; then foreign outsourcing and, for the highest productivity firms, FDI. An additional fixed cost is incurred for commencing operations in a new country; thus one observes a positive correlation between firm productivity and the number of trade and investment destinations and sources.⁴⁶ Firms engaged in foreign production are not limited to importing back to the home country; they can also sell in the host country or in third countries.

Form of relationship	Location of production	
	Home	Abroad
Make: Vertical Integration	no trade	FDI & intra-firm trade
Buy: Outsource	no trade	import arm's length

⁴⁴ Specifically, measures of innovation are approximately quadratic in the Lerner index; the Lerner index ranges from zero (perfect competition) to one (perfect monopoly).

⁴⁵ The key papers in this literature are Antràs (2003), Antràs and Helpman (2004 and 2008), and Helpman *et al.* (2004); the studies reviewed in Görg and Greenaway predate these theoretical innovations. Keller (2010) reviews the subsequent literature. Grossman and Rossi-Hansberg (2008) advance a theory of production fragmentation based on trade in tasks.

⁴⁶ See Eaton *et al.* (2011) for these patterns in France; De Hoyos and Iacovone (2011) find this pattern for Mexican firms.



These theoretical developments drawing on organization theory have yet to be tested with micro-level data; this is the direction of current empirical analysis in this area. Like other strands of heterogeneous-firm trade theory, its firm-level focus challenges international trade theories constructed on the nation state as the fundamental unit of analysis.

Conclusion Part II

The discussion of the Melitz model noted that Pareto distributions

accurately describe the distribution of firms by size and by productivity and that simulating changes in these distributions is one approach emerging in the current literature. Eaton *et al.* (2011) provides an illustration of this simulation work. The study uses French firm-level data to simulate the impact of a 10-percent uniform reduction in trade costs on the distribution of trading and non-trading firms by size (firm size and firm productivity are highly correlated). The graph below plots the selection impact of liberalization by firm-size decile (all firms, not just trading firms). Half of the firms in the lowest decile exit; there is net exit in all deciles. The impact on firm sales is more pronounced, sales decline for all but the top decile.

Selection, the culling of smaller, less-efficient firms, and the reallocation of factors to larger, more productive firms is a major source of productivity growth. This is what Griliches (1963) identified as a source of aggregate productivity growth; a source of growth excluded by assumption in aggregate growth models. It is impressive that after 50 years the economics profession has access to the data and has developed models that are beginning to explain aggregate productivity growth from the micro level.

The R&D spillover theories discussed in Part I are compelling narratives if one assumes universal technical efficiency at the firm level. If all firms operate on the efficiency frontier, then the only possible source of TFP growth is a positive shift in the frontier: that is, all observed growth is the direct result of technical innovation generated by R&D. But the empirical evidence review in Part II does not support the distribution of technical efficiency assumed in exogenous and endogenous growth models. It finds that most firms operate far inside the efficiency frontier: X-efficiency is pervasive. Most of what is observed as TFP growth at the national and industrial level is selection-driven and involves movements toward the technology frontier, rather than movements with the frontier. The frontier does shift because of R&D-driven innovation, but this is only one of several factors that influence aggregate productivity.

PART III: CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH AND AGRICULTURAL TRADE

The objective of this paper is to provide a critical literature review as background to applied research on the benefits of agricultural trade to firms in developing countries, and specifically on trade-related R&D spillovers. Part I of the paper reviews the literature on R&D spillovers at the macro-level, where the nation economy is the unit of analysis. Part II reviews several emerging lines of micro-level research, where the firm is the unit of analysis, that focus on the relationship between international trade and firm-level and industry-level productivity. Both parts share a focus on the relationship between trade and productivity growth but the causal channels the two bodies of research identify differ fundamentally. R&D is central to the macro-level literature reviewed in Part I. In micro-level studies R&D is at best peripheral: selection and learning are the primary causal channels between trade and productivity at the firm level.

The difference between the macro and micro literatures follows directly from differences in the assumptions and definitions employed. The macro-level analyses are part of the endogenous economic growth literature. In this framework economic growth can be attributed to two causes: 1) increases in factors and 2) technological change, which is the direct product of R&D. Thus, total factor productivity growth (that is, growth net of growth in productive factors) is determined by R&D. International trade is hypothesized to increase the effectiveness of national-level R&D by increasing the likelihood of R&D spillovers from trading partners. An international R&D spillover is narrowly defined: it is the flow of knowledge from one country to another that leads to the production of new knowledge in the recipient country. There is confusion in the literature about R&D spillovers, as Zvi Griliches has noted.⁴⁷

[T]here are two distinct notions of R&D “spillovers” here which are often confused in the literature. In the first, R&D intensive inputs are purchased from other industries at less than their full “quality” price. ... If capital equipment purchase price indices reflected fully the improvements in their quality, i.e., were based on hedonic calculations, there would be no need to deal with it. As currently measured, however, total factor productivity in industry *i* is affected not only by its own R&D but also by productivity improvements in industry *j* to the extent of its purchases from that industry and to the extent that the improvements in *j* have not been appropriated by its producers and/or have not been incorporated in the official price indices of that (*i*) industry by the relevant statistical agencies. The use of purchase-flow-weighted R&D measures assumes that social returns in industry *j* are proportional to its R&D investment levels and that the amount of such returns transferred to industry *i* is proportional to its purchases from industry *j*.

But these are not real knowledge spillovers. They are just consequences of conventional measurement problems. True spillovers are the ideas borrowed by the research teams of industry *i* from the research results of industry *j*. It is not clear that this kind of borrowing is particularly related to input purchase flows.⁴⁸

The confusion between the two notions of R&D spillover noted by Griliches has been common in the literature measuring the social returns to public agricultural R&D. The assumption is that there is a

⁴⁷ This quotation from Griliches 1992 is an almost verbatim repetition of Griliches 1979 (1998: 30-31): the confusion in the literature is now in its fourth decade.

⁴⁸ Griliches 1998 [1992]: 257-58.

compelling market failure in R&D for agriculture. Because it is not economically rational for individual farmers to undertake R&D and because the private returns from agricultural R&D are assumed to be largely unappropriable, private agents will under-invest in agricultural R&D. This is rationale for public provision of agriculture R&D. To justify this use of public funds agricultural economists at public institutions estimate the social benefits and the social rate of return of these public investments. In calculating the increase in producers' and consumers' surplus from the adoption and diffusion of, for example, an improved variety developed by a public research station, one does not quality-adjust the new variety: it is the social value of the publically-financed quality improvement that one is attempting to measure. When new varieties diffuse across state or national borders a spillover exists. Such spillovers were often referred to as R&D spillovers; however, such spillovers are not pure knowledge spillovers as defined by Griliches and as used in endogenous growth theory. They are, in fact, unappropriated quality improvements in inputs adopted by private agents in the agricultural sector. To eliminate confusion, in light of endogenous growth theory, agricultural economists now refrain from using "R&D spillover": they use the term "technology spillover" to refer to the cross-border diffusion and adoption of inputs embodying R&D; and they use "knowledge spillover" to refer to pure knowledge spillovers.

With this distinction in mind one can restate the conclusion of Part I: the empirical approach of the macro literature, regressing national TFP growth on various international purchase-flow-weighted R&D measures, conflates knowledge spillovers and technology spillovers. The finding that capital-good-import-weighted R&D generates the best regression coefficients is consistent with the insufficient quality adjustment of imported capital goods. It is likely that some or much of what is being measured is unappropriated productivity improvements by foreign suppliers of capital goods and other inputs. This problem is to be expected when an indirect indicator of innovation derived from aggregate secondary data (TFP growth) is used. Credible evidence of international knowledge spillovers requires direct or more proximate observation of institutions engaged in R&D; if this is an important hypothesis then funding for primary data collection should become a priority.

The micro-level research surveyed in Part II is more inductive and exploratory than theoretically-derived. The underlying common denominator is an attempt to identify sources of the dynamic gains from trade. Dynamic in this context is opposed to static: dynamic gains are those in excess of the static gains. In terms of the standard welfare-gains-from-trade diagram, static gains result from movements along fixed domestic supply and demand curves; dynamic gains follow from movements in the curves. Given the firm-level panel data available, the focus is on the sources of productivity gains on the supply side; on whether, how and to what extent the domestic supply curve shifts rightward. The dominant importance of selection effects, especially the reallocation of productive factors from less productive exiting firms to more productive surviving firms, is consistent with a rightward/downward shift of the domestic supply curve. Learning effects are also consistent with dynamic gains. Thus a consistent, empirically-grounded narrative emerges: increased international competition induces selection and productivity-improving investment, pre-trade and sometimes post-trade. The panel data sets that provide the empirical base do not provide evidence of international spillovers, whether of technical spillovers or knowledge spillovers. That many of the panel-data based studies use labour productivity rather than total factor productivity leaves scope for technical spillovers. It is unlikely that knowledge spillovers play a major role in the productivity gains observed in the panel studies, given the low proportion of firms that engages in R&D. Focusing a study on firms that do engage in R&D would allow one to gauge the relative magnitude of the international knowledge spillovers.

The key question facing the FAO project on analyzing the benefits of agriculture trade to developing countries based on firm and industry behavior is whether to restrict its empirical investigation into

international R&D spillovers to knowledge spillovers. The propensity of agro-industrial firms in developing countries to engage in R&D is even less than in developed countries. This limits the likely importance of knowledge spillovers as a benefit of trade, but the small number of R&D-engaged firms in any given developing country may make comprehensive in-depth surveys and interviews feasible. The alternatives are to broaden the scope slightly to allow technical spillovers or broaden the scope even further to allow the full range of causal channels under the heading of dynamic gains from trade.

Invariant of the choice of these three alternatives, there are some methodological lessons that can be drawn from empirical analyses reviewed in parts I and II. First, adopt a micro-level, firm-based focus; the weaknesses of macro-level analysis have been noted many times in this paper. Second, collect primary data. Most of the studies reviewed in Part II are based on panel data sets constructed by national governments or multi-lateral development institutions. The cost and years involved makes a panel infeasible; but another panel is not necessary. The stylized facts are well-established: there are selection effects and learning effects and these can be examined in a small, well-designed sample that matches trading and non-trading firms, for example. Third and most important, what is missing from almost all of the studies reviewed in this paper is qualitative data. The quantitative results indicate that firms do things that make them more productive prior to engaging in trade and sometimes once they are engaged in trading: but this is merely statistical inference. Few studies engage firms directly and ask executives and managers what they did, why they did it, when they did it, and what they might have done had the institutional environment been different. The answers to questions like these, which cannot be posed to a secondary data set, have the potential to move the research program forward and inform and refine future research.

Van Biesebroeck (2005) is one of the few studies to find strong and significant evidence of post-entry productivity gains by exporters. The study is of a panel of manufacturing firms in nine Sub-Saharan countries. The study is exceptional in that it uses qualitative data in addition to quantitative data. Van Biesebroeck's argument, based on a (qualitative) survey of panel firm executives, is that contract enforcement in the home market is weak. In contrast, the risk of non-payment by foreign customers is very low. Thus new exporters gain increased access to credit and realize in the export market economies of scale that had been limited by a lack of reliable domestic customers. This is a post-entry, export-induced productivity effect but it is neither a knowledge (R&D) spillover nor a technology spillover nor does it qualify as a learning effect. It is a causal path that one could not have inferred from quantitative results alone. Such qualitative findings enrich the emerging narrative about the variety of benefits of trade to developing countries and about local impediments to development; such findings are important because they inform the direction of data collection, theory construction and empirical analysis.

APPENDIX A: Measuring stocks of knowledge

The stock of knowledge (R&D) is constructed from periodic (usually annual) R&D expenditure data, a flow value, using the Perpetual Inventory Method (PIM).⁴⁹ The R&D stock (S_t) at the end of period t is equal to the beginning stock (S_{t-1}) plus R&D expenditure during the year (R_t), minus depreciation of the beginning stock (δS_{t-1}), where δ is the annual depreciation rate.

$$S_t = (1-\delta)S_{t-1} + R_t$$

The stock of R&D in the initial year (S_0) is constructed thus: $S_0 = R_1 / (\delta + g)$; where g is the average annual logarithmic growth rate of R&D from the initial year to the present. The initial stock is not observed, it is constructed based on the assumption that R&D spending and depreciation prior the initial period is the same as the average rates after the initial period. The absolute stock of R&D is sensitive to the validity of this assumption.

The depreciation rate is generally assumed to be 5%; this means it takes 13.5 years for a given stock to depreciate by half and 45 years to depreciate by 90%. If one takes endogenous growth theory seriously then the depreciation rate should be zero for the horizontal (Dixit-Stiglitz product variety) model: knowledge once created is assumed to be immortal and demand for increasing variety literally implies that no product becomes obsolete. The depreciation rate for the vertical model, which assumes continual creative destruction, rendering existing stocks of knowledge obsolete, should logically have a relatively high rate of depreciation; it would likely vary considerably year-to-year as well.⁵⁰

Agricultural economists have a distinct approach to constructing agricultural R&D stocks. The convention is to assume agricultural research (e.g., plant breeding) takes several years to before any useful research product becomes available. The new product then needs to be tested and, if viable, scaled up. Thus R&D is lagged several years, its effective value increases gradually, reaches a peak or plateau in its mature phase, and then becomes obsolete as new, improved varieties are released or as it ceases to be resistant to pests and diseases. The time profile of agricultural R&D is usually a trapezoid or a gamma-distribution density function.⁵¹

⁴⁹ The OECD MSTI (Main Science and Technology Indicators) database is the standard source. The methodology of the indicators is elaborated in the *Methodology of the Frascati Manual*, OECD (2002a).

⁵⁰ Coe *et al.* (2009), for example, assume a horizontal model and use 5 percent depreciation. They test whether their results are sensitive to this assumption and find that they are not (they test $\delta = .0$ and $.2$). This further supports the argument presented in the conclusion of part I about the low precision of R&D data.

⁵¹ There is a huge literature on this topic; Sheng *et al.* (2011) has an excellent applied discussion of the construction R&D (knowledge) stocks for agricultural research and a good bibliography; it also examines R&D spillovers into Australian agriculture. Adams (1990) provides empirical evidence of long (15-20 year) lags for basic research. Pakes and Schankerman (1984) explore obsolescence and gestation lags.

APPENDIX B : Grossman and Helpman on weighting international information flows

The passage below is from Grossman and Helpman (1991) *Innovation and Growth in the Global Economy*; pp. 166-67.

It is plausible to suppose that the foreign contribution to the local knowledge stock increases with the number of commercial interactions between domestic and foreign agents. That is, we may assume that international trade in tangible commodities facilitates the exchange of intangible ideas. This assumption can be justified in several ways. First, the larger the volume of international trade, the greater presumably will be the number of personal contacts between domestic and foreign individuals. These contacts may give rise to an exchange of information and may cause the agents from the small country to acquire novel (for them) perspectives on technical problems. Second, imports may embody differentiated intermediates that are not available in the local economy. The greater the quantity of such imports, the greater perhaps will be the number of insights that local researchers gain from inspecting and using these goods. Third, when local goods are exported, the foreign purchasing agents may suggest ways to improve the manufacturing process. In the context of our model, the recommendations might take the form of ideas for new intermediate inputs. The number of such suggestions is likely to increase with the quantity of goods exported. It seems reasonable to assume therefore that the extent of the spillovers between any two countries increases with the volume of their bilateral trade.

To pursue the implications of this hypothesis, we let $K_n(t)$ denote the stock of knowledge capital in the small country, and suppose that the growth of K_n depends not only on spillovers from local research but also on those international contacts. In particular, we specify $K_n(t) = G[n(t), T(t)]$, where T represents the cumulative volume of trade (exports plus imports) up to time t .

BIBLIOGRAPHY

- Acemoglu, D., Aghion, P. & Zilibotti, F. 2006. Distance to frontier, selection and economic growth. *Journal of the European Economic Association*, 4(1): 37-74.
- Acemoglu, D., Antràs, P. & Helpman, E. 2007. Contracts and technology adoption. *American Economic Review*, 97(3): 916-943.
- Acemoglu, D., Johnson, S. & Robinson, J.A. 2001. The colonial origins of comparative development: an empirical investigation. *American Economic Review*, 91(5): 1369-1401.
- Acemoglu, D., Johnson, S. & Robinson, J.A. 2005. Institutions as a fundamental cause of long-run growth. In P. Aghion & S. Durlauf, eds. *Handbook of Economic Growth, Volume 1A*, pp. 385-472. Amsterdam, NL, Elsevier.
- Acharya, R.C. & Keller, W. 2008. Estimating the productivity selection and spillover effects of imports. NBER working paper 14079. Cambridge, MA, USA, NBER Press.
- Adams, J.D. 1990. Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, 98(4): 673-702.
- Adelman, I.G. 1958. A stochastic analysis of the size distribution of firms. *Journal of the American Statistical Association*, 53(284): 898-904.
- Aghion, P., Bloom, N., Blundell, R., Griffith, R. & Howitt, P. 2005. Competition and innovation: an inverted-u relationship. *Quarterly Journal of Economics*, 120 (2): 701–728.
- Aghion, P. & Howitt, P. 1992. A model of growth through creative destruction. *Econometrica*, 60(2): 323–51.
- Alston, J.M. 2002. Spillovers. *Australian Journal of Agricultural and Resource Economics*, 46(3): 315-46.
- Amiti, M. & Konings, J. 2007. Trade liberalization, intermediate inputs, and productivity. *American Economic Review*, 97(5): 1611-38.
- Amiti, M. & Khandelwal, A.K. 2009. Import competition and quality upgrading. NBER working paper 15503. Cambridge, MA, USA, NBER Press.
- Arnold, J.M. & Hussinger, K. 2005. Export behavior and firm productivity in German manufacturing: a firm-level analysis. *Review of World Economics*, 141: 219-243.
- Antràs, P. 2003. Firms, contracts, and trade structure. *Quarterly Journal of Economics*, 118(4): 1375-1418.
- Antràs, P. & Helpman, E. 2004. Global sourcing, *Journal of Political Economy*, 112(3): 552-80.
- Antràs, P. & Helpman, E. 2008. Contractual frictions and global sourcing. In E. Helpman, D. Marin & T. Verdier, eds. *The Organization of Firms in a Global Economy*, pp. 9-54. Cambridge, MA, USA, Harvard University Press.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. In R. Nelson, ed. *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609-625. Princeton, NJ, Princeton University Press.

Asplund, M. & Nocke, V. 2006. Firm turnover in imperfectly competitive markets. *Review of Economic Studies*, 73(2): 295-327.

Audretsch, D. & Feldman, M. 2004. Knowledge spillovers and the geography of innovation. In: J.V. Henderson and J.F. Thisse, (eds), *Handbook of Regional and Urban Economics*, Volume IV, pp. 2714-2739. Amsterdam, NL, North-Holland,

Aw, B.Y., Chen, X. & Roberts, M.J. 2001. Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing. *Journal of Development Economics*, 66(1): 51-86.

Aw, B.Y., Roberts, M.J. & Xu, D.Y. 2008. R&D investments, exporting, and the evolution of firm productivity. *American Economic Review*, 98(2): 451-456.

Aw, B.Y., Roberts, M.J. & Xu, D.Y. 2011. R&D investment, exporting and productivity dynamics, *American Economic Review*, 101(4): 1312-1344.

Baily, M. N., Hulten, C. & Campbell, D. 1992. Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity, Microeconomics*, 1992: 187-249.

Baltagi, B.H. 2008. *Econometric Analysis of Panel Data*, 4th edition. Chichester, UK, John Wiley & Sons.

Barbosa, N. & Faria, A.P. 2011. Innovation across Europe: how important are institutional differences? *Research Policy*, 40(9): 1157-1169.

Bartelsman, E. J., & Doms, M. 2000. Understanding productivity: lessons from longitudinal microdata. *Journal of Economic Literature*, 38 (3), 569-594.

Bartelsman, E. J., Haskel, J. & Martin, R. 2008. Distance to which frontier? Evidence on productivity convergence from international firm-level data. *CEPR Discussion Paper 7032*.

Benhabib, J. & Spiegel, M. 1995. The role of human capital in economic development: evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34: 143-173

Benhabib, J. & Spiegel, M. 2005. Human capital and technological diffusion. In P. Aghion & S. Durlauf, eds. *Handbook of Economic Growth*, Volume 1A, pp. 935-966. Amsterdam, NL, Elsevier.

Bernard A.B. & Jensen, J.B. 1995. Exporters, jobs, and wages in US manufacturing: 1976-87. *Brookings Papers on Economic Activity: Microeconomics*, 1995: 67-112.

Bernard, A. B. & Jensen, J.B. 1999. Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics*, 47 (1): 1-25.

Bernard, A. B. & Jensen, J. B. 2004. Why some firms export. *Review of Economics and Statistics*, 86(2): 561-569.

Bernard A.B., Eaton, J., Jensen, J.B. & Kortum, S.S. 2003. Plants and productivity in international trade. *American Economic Review*, 93(4): 1268-90.

Bernard, A.B., Jensen, J.B., Redding, S. & Schott, P. 2007. Firms in international trade. *Journal of Economic Perspectives*, 21(3):105-130.

Bernard A.B., Jensen, J.B. & Schott, P.K. 2009. Importers, exporters and multinationals: a portrait of firms in the U.S. that trade goods. In T.Dunne, J.B. Jensen & M.J Roberts, eds. *Producer Dynamics: New Evidence from Micro Data*, pp. xxxxxx. Chicago, IL, USA, University of Chicago Press.

Bernard, A.B., Jensen, J.B. & Schott, P.K. 2006. Trade costs, firms and productivity. *Journal of Monetary Economics*, 53(5): 917-937.

Bernard, A.B., Jensen, J.B., Redding, S.J & Schott, P.K. 2011. The empirics of firm heterogeneity and international trade. NBER working paper 17627. Cambridge, MA, USA, NBER Press.

Bloom, N. & Van Reenen, J. 2007. Measuring and explaining management practices across firms and countries, *Quarterly Journal of Economics* 122(4): 1351-1408.

Bloom, N., Sadun, R. & Van Reenen, J. 2010a. Recent advances in the empirics of organization economics. *Annual Review of Economics*, 2: 105-37.

Bloom, N., Propper, C., Seiler, S. & Van Reenen, J. 2010b. The impact of competition on management quality: evidence from public hospitals. NBER working paper no. 16032.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. & Roberts, J. 2011a. Does management matter? Evidence from India. NBER working paper no. 16658. Cambridge, MA, USA, NBER Press.

Bloom, N., Draca, M., & Van Reenen, J. 2011b. Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. NBER working paper no. 16717. Cambridge, MA, USA, NBER Press.

Botero, J. C., Djankov, S., La Porta, R., Lopez-de-Silanes, F. & Shleifer, A. 2004. The regulation of labour. *Quarterly Journal of Economics*, 119(5): 1339-1382.

Braguinsky, S., Branstetter, L. & Regaterio, A. 2011. The incredible shrinking Portuguese firm. NBER Working Paper No. 17265. Cambridge, MA, USA, NBER Press.

Bustos, P. 2011. Trade liberalization, exports and technology upgrading: evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review*, 101(1): 304-40.

Caves, R.E. 1998. Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, 36(4): 1947-1982.

Chamberlin, E.H. 1933. *The Theory of Monopolistic Competition*. Cambridge, MA, USA, Harvard University Press.

Chaney, T. 2008. Distorted Gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98(4): 1707-21.

- Chaney, T. 2011. The network structure of international trade. NBER working paper 16753. Cambridge, MA, USA, NBER Press.
- Clerides, S., Lach, S. & Tybout, J. 1998. Is Learning-by-exporting important? Micro-dynamic evidence from Colombia, Morocco, and Mexico. *Quarterly Journal of Economics*, 113(3): 903-47
- Coe, D.T. & Helpman, E. 1995. International R&D spillovers. *European Economic Review*, 39(5): 859-887.
- Coe, D.T., Helpman, E. & Hoffmaister, A. 1997. North-South R&D spillovers. *Economic Journal*, 107(440): 134-49.
- Coe, D. T., Helpman, E. & Hoffmaister, A.W. 2009. International R&D spillovers and institutions. *European Economic Review*, 53(5): 723–741.
- Cohen, D. & Soto, M. 2007. Growth and human capital: good data, good results. *Journal of Economic Growth*, 12(1): 51-76.
- Cohen, W. M. & S. Klepper (1992) The anatomy of industry R&D intensity distributions. *American Economic Review*, 82(4): 773–88.
- Cohen, W.M. & Levinthal, D.A., 1989. Innovation and learning: the two faces of R&D. *Economic Journal*, 99(4): 569-596.
- Conley, T. G., & Udry, C.R. 2010. Learning about a new technology: pineapple in Ghana. *American Economic Review*, 100(1): 35–69
- Costantini, J.A. & Melitz, M.J. 2008. The dynamics of firm-level adjustment to trade. In E. Helpman, D. Marin & T. Verdier, eds. *The Organization of Firms in a Global Economy*, pp. 107-142. Cambridge, MA, USA, Harvard University Press.
- Damijan, J.P., Kostevc, C. & Polanec, S. 2008. From innovation to exporting or vice versa? Causal link between innovation activity and exporting in Slovenian microdata. Leuven, BE, LICOS discussion paper 20408.
- Dasgupta, P. & Stiglitz, J.E. 1980a. Uncertainty, industrial structure, and the speed of R&D. *Bell Journal of Economics*, 11(1): 1-28.
- Dasgupta, P. & Stiglitz, J.E. 1980b. Industrial structure and the nature of innovative activity. *Economic Journal*, 90(358): 266-293.
- De Hoyos, R. E. & Iacovone, L. 2011. Economic performance under NAFTA : a firm-level analysis of the trade-productivity linkages. Washington, DC, World Bank Policy Research Working Paper Series no. 5661
- Davis, S. J. & Haltiwanger, J. 1999. Gross job flows. In O. Ashenfelter & D. Card, eds. *Handbook of Labour Economics*, vol. 3B, pp. 2711–2805. New York, NY, USA, Elsevier Science.
- Dixit, A.K. & Stiglitz, J.E. 1977. Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3): 297–308.

Dunne, T., Roberts, M. & Samuelson, L. 1989. The growth and failure of US manufacturing plants. *Quarterly Journal of Economics*, 104(4): 671-698.

Durlauf, S.N, Johnson, P. & Temple, P. 2005. Growth econometrics. In P. Aghion & S. Durlauf, eds. *Handbook of Economic Growth* vol 1A, pp. 555-677. Amsterdam, NL, Elsevier.

Eaton, J., Kortum, S. & Kramarz, F. 2011. An anatomy of international trade: evidence from French firms. *Econometrica*, 79(5): 1453-1498.

Engelbrecht, H.-J. 1997. International R&D spillovers, human capital and productivity in OECD economies: an empirical investigation. *European Economic Review*, 41(8): 1479-1488.

Eslava, M., Haltiwanger, J.C., Kugler, A.D. & Kugler, M. 2009. Trade reforms and market selection: evidence from manufacturing plants in Colombia. NBER working paper 14935. Cambridge, MA, USA, NBER Press.

Ethier, W. J. 1982. National and international returns to scale in the modern theory of international trade. *American Economic Review*, 72(3): 389-405.

Ethier, W.J. 2004. Dixit-Stiglitz, trade and growth. In S. Brakman & B.J. Heijdra, eds.: *The Monopolistic Competition Revolution in Retrospect*, pp. 149-156. New York, NY, USA, Cambridge University Press.

Fernandes, A. M. & Isgut, A.E. 2005. Learning-by-doing, learning-by-exporting, and productivity: evidence from Colombia. Washington, DC, World Bank Policy Research Working Paper Series no. 3544.

Foster, L., Haltiwanger, J. & Krizan, C.J. 2001. Aggregate productivity growth: lessons from microeconomic evidence, in E. Dean, M. Harper & C. Hulten, eds. *New Developments in Productivity Analysis*, pp. 303-363. Chicago, IL, USA, University of Chicago Press.

Foster, L., Haltiwanger, J. & Krizan, C.J. 2006. Market selection, reallocation, and restructuring in the U.S. retail trade sector in the 1990s. *Review of Economics and Statistics*, 88(4): 748-58.

Foster, L., Haltiwanger, J. & Syverson, C. 2008. Reallocation, firm turnover, and efficiency: selection on productivity or profitability? *American Economic Review*, 98(1): 394-425.

Fujita, M. & Thisse, J.-F. 2002. *Economics of Agglomeration. Cities, Industrial Location and Regional Growth*. New York, NY, USA, Cambridge University Press.

Funk, M. 2001. Trade and international R&D spillovers among OECD countries. *Southern Economic Journal*, 67(3): 725-736.

Girma, S., Greenaway, D. & Kneller, R. 2004. Does exporting increase productivity? A microeconomic analysis of matched firms. *Review of International Economics*, 12(5): 855-66.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A. & Shleifer, A. 1992. Growth in cities. *Journal of Political Economy*, 100(6): 1126-52.

Görg, H. & Greenaway, D. 2004. Much ado about nothing? Do domestic firms really benefit from foreign direct investment? *World Bank Research Observer*, 19(2): 171-197.

Greenaway, D. & Kneller, R.A. 2008. Exporting, productivity and agglomeration. *European Economic Review*, 52(5): 919-939.

Griliches, Z. 1957. Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4): 501-22.

Griliches, Z. 1979. Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 10 (1): 92–116.

Griliches, Z., ed. 1984. *R&D, Patents and Productivity*. Chicago, IL, USA, University of Chicago Press.

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

Griliches, Z. & Ringstad, V. 1971. *Economies of Scale and the Form of the Production Function*. Amsterdam, NL, North-Holland

Grossman, G. M. & Helpman, E. 1991a. *Innovation and Growth in the Global Economy*. Cambridge, MA, USA, MIT Press.

Grossman, G. M. & Helpman, E. 1991b. Quality ladders in the theory of growth. *Review of Economic Studies*, 58(1): 43-61.

Grossman G.M. & Rossi-Hansberg, E. 2008. Trading tasks: a simple theory of offshoring. *American Economic Review*, 98(5): 1978-97.

Gustafsson, P. & Segerstrom, P. 2010. Trade liberalization and productivity growth. *Review of International Economics*, 18(2): 207-228.

Helpman, E., ed. 2008. *Institutions and Economic Performance*. Cambridge, MA, USA, Harvard University Press.

Helpman E, Melitz, M.J. & Rubinstein, Y. 2008. Estimating trade flows: trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2): 441-87.

Helpman E, Melitz, M.J. & Yeaple, S.R. 2004. Export versus FDI with heterogeneous firms. *American Economic Review*, 94(1): 300-16.

Hoch, I. 1955. Estimation of production function parameters and testing for efficiency. *Econometrica*, 23(3): 325-326.

Hoch, I. 1958. Simultaneous equations bias in the context of the Cobb-Douglas production function. *Econometrica*, 26(1): 566-578.

Hoch, I. 1962. Estimation of production function parameters combining time-series and cross-sectional data. *Econometrica*, 30(1): 34-53.

Hopenhayn, H. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60 (5): 1127–50.

Hoppe, R.A., MacDonald, J.M. & Korb, P. 2010. Small farms in the United States: persistence under pressure, EIB 63. Washington, DC, USDA, Economic Research Service.

Hsiao, C. 2003. Analysis of Panel Data, 2nd edition. New York, NY, USA, Cambridge University Press.

Hsieh, C-T. & Klenow, P. 2009. Misallocation and manufacturing TFP in China and India. Quarterly Journal of Economics, 124(4):1403–48.

Iacovone, L., Keller, W. & Rauch, F. 2011. Innovation responses to import competition, Seminar paper, Department of Economics, Princeton University.

ISGEP. 2008. Understanding cross-country differences in exporter premia: comparable evidence for 14 countries. Review of World Economics, 144(4): 596-635.

Jaffe, A.B. 1986. Technological opportunity and spillovers of R&D: evidence from firms patents, profits, and market value. American Economic Review, 76(5): 984–1001.

Jaffe, A. B., Trajtenberg, M. & Henderson, R.M. 1993. Geographic localization of knowledge spillovers as evidenced by patent citation. Quarterly Journal of Economics, 108(3): 577- 598.

Jones, C. 1995. Time series tests of endogenous growth models. Quarterly Journal of Economics, 110(2): 495-525.

Jones, C. 2005. Growth and ideas. In P. Aghion & S. Durlauf, eds. Handbook of Economic Growth Vol 1B, pp. 1063-1111. Amsterdam, NL, Elsevier.

Jorgenson, D.W. & Griliches, Z. 1967. The explanation of productivity change. Review of Economic Studies, 4(3): 249-83.

Jorgenson, D.W. & Griliches, Z. 1972. Issues in growth accounting: A Reply to Edward F. Denison. Survey of Current Business, 52(5-Part II): 65-94.

Jorgenson, D.W. & Vu, K.M. 2011. The rise of developing Asia and the new economic order. Journal of Policy Modeling, 33: 698–745.

Jovanovic, B. (1982) Selection and the evolution of industry, Econometrica 50 (3): 649–70.

Jovanovic, B. & Rob, R. 1989. The growth and diffusion of knowledge. Review of Economic Studies, 56 (4): 569–82.

Keller, W. 1998. Are international R&D spillovers trade-related? Analyzing spillovers among randomly matched trade partners. European Economic Review, 42(8): 1469-1481.

Keller, W. 2004. International technology diffusion. Journal of Economic Literature, 42(3): 752-782.

Keller, W. 2010. International trade, foreign direct investment, and technology spillovers. In B.H. Hall & N. Rosenberg, eds. Handbook of the Economics of Innovation, pp. 793-829. Amsterdam, NL, Elsevier.

Krugman, P.R. 1979. Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9(4): 469-479.

Krugman, P.R. 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70(5): 950-959.

Krugman, P.R. 1991. Increasing returns and economic geography. *Journal of Political Economy*, 99(3): 483-499.

La Porta, R., Lopez-de-Silanes, F. & Shleifer, A. 1999. Corporate ownership around the world. *Journal of Finance*, 54 (2): 471-517.

Leibenstein, H. 1966. Allocative efficiency vs. X-efficiency. *American Economic Review*, 56 (3): 392-415

Levin, R. & Reiss, P. 1984. Tests of a Schumpeterian model of R&D and market structure. In

Z. Griliches, ed. *R&D, Patents and Productivity*, pp. 175-208. Chicago, IL, USA, University of Chicago Press.

Levin, R. & Reiss, P. 1988. Cost-reducing and demand-creating R&D with spillovers. *Rand*

Journal of Economics, 19(4): 538-56.

Levin, R.C., Klevorick, A., Nelson, R.R. & Winter, S.C. 1987. Appropriating the returns from industrial R&D. *Brookings Papers on Economic Activity*, 1987(3): 783-820.

Levinsohn, J. & Petrin, A. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2): 317-41.

Lichtenberg, F. & van Pottelsberghe de la Potterie, B. 1998. International R&D spillovers: a comment. *European Economic Review*, 42(8): 1483-1491.

Lileeva, A. & Trefler, D. 2010. Improved access to foreign markets raises plant-level productivity . . . for some plants. *Quarterly Journal of Economics*, 125(3): 1051-99.

Liu, L. & Tybout, J.R. 1996. Productivity growth in Chile and Columbia: the role of entry, exit and learning. In M. J. Roberts & J. R. Tybout, eds. *Industrial evolution in developing countries: Micro patterns of turnover, productivity and market structure*, pp. 73-103. New York, NY, USA, Oxford University Press.

López, R. A. 2005. Trade and growth: reconciling the macroeconomic and microeconomic evidence. *Journal of Economic Surveys*, 19(4): 623-648.

López, R. A. 2009. Do firms increase productivity in order to become exporters? *Oxford Bulletin of Economics and Statistics*, 71(5): 621-642.

López, R. A. & Yadav, N. 2010 Imports of intermediate inputs and spillover effects: evidence from Chilean plants. *Journal of Development Studies*, 46 (8): 1385-1403.

Lucas, R.E. 1978. On the size distribution of business firms. *Bell Journal of Economics*, 9(2): 508-523.

- Lucas, R. E. 1988. On the mechanics of economic development. *Journal of Monetary Economics*, 22 (1): 3-42.
- Mansfield, E. 1962. Entry, Gibrat's Law, innovation, and the growth of firms. *American Economic Review*, 52(5): 1023-51.
- Mansfield, E., Rapoport, J., Romeo, A., Wagner, S. & Beardsley, G. 1977. Social and private rates of return from industrial innovations. *Quarterly Journal of Economics*, 91(2): 221-240.
- Marschak, J. & Andrews, W.H. 1944. Random simultaneous equations and the theory of production. *Econometrica*, 12 (1): 143-205.
- Mundlak, Y. 1961. Empirical production functions free of management bias. *Journal of Farm Economics*, 43(1): 44-56.
- Nasbeth, L. & Ray, G. eds. 1974. *The Diffusion of New Industrial Processes: an International Study*. Cambridge, UK, Cambridge University Press.
- Nelson, R. & Phelps, E. 1966 Investment in humans, technological diffusion, and economic growth. *American Economic Review*, 56(1): 69-75.
- Nerlove, M. 2002. The history of panel data econometrics, 1961-1997. In M.Nerlove, ed. *Essays in Panel Data Econometrics*, pp. 1-114. Cambridge, UK, Cambridge University Press.
- Nickell, S. 1996. Competition and corporate performance. *Journal of Political Economy*, 104(4): 724-746.
- Olley, G. S. & Pakes, A. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64 (6): 1263-97.
- OECD. 2002a. *Methodology of the Frascati Manual*. Paris, OECD. (also available at www.oecd.org/sti/frascaticmanual).
- OECD. 2002b. *Productivity and innovation: the impact of product and labour market policies*. OECD Economic Outlook 71.
- Pakes, A. & Schankerman, M.A. 1984. The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. In Z. Griliches, ed. *R&D, Patents, and Productivity*, pp. 209-232. Chicago, IL, USA, University of Chicago Press.
- Park, A., Yang, D., Shi, X & Jiang, Y. 2010. Exporting and firm performance: Chinese exporters and the Asian financial crisis. *Review of Economics and Statistics*, 92(4): 822-842.
- Pavcnik, N. 2002. Trade liberalization, exit, and productivity improvement: evidence from Chilean plants. *Review of Economic Studies*, 69(1): 245-76.
- Redding, S.J. 2010. *Theories of heterogeneous firms and trade*. NBER working paper 16562, Cambridge, MA, USA, NBER Press.
- Roberts, M. & Tybout, J. 1997. Producer turnover and productivity growth in developing countries. *World Bank Research Observer*, 12(1): 1-18.

Rogers, E.M & Beal, G.M. 1958. Reference group influences in the adoption of agricultural technology. Report 1236, Ames, IA, USA, Iowa Agricultural and Home Economics Experiment Station.

Rogers, E.M. 1962. Diffusion of Innovations, Glencoe, IL, USA, Free Press.

Romer, P. M. 1986. Increasing returns and long-run growth. *Journal of Political Economy*, 94 (5): 1002-1037.

Romer, P.M. 1990. Endogenous technical change. *Journal of Political Economy*, 98(5-Part 2): S71-S102.

Ryan, B. & Gross, N.C. 1943. The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology*, 8: 15–24.

Ryan, B. & Gross, N.C. 1950. Acceptance and diffusion of hybrid corn seed in two Iowa communities. Research Bulletin 372, Ames, Iowa, USA, Iowa Agricultural Experiment Station.

Saint-Paul, G. 2002. Employment protection, international specialisation and innovation. *European Economic Review*, 46(2): 375-395.

Schreyer, P. 2001. OECD Manual: Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth. Paris: OECD.

Schreyer, P. 2002. Computer price indices and international growth and productivity comparisons. *Review of Income and Wealth*, 48: 15–31.

Schreyer, P. 2009. OECD Manual: Measuring Capital. Paris: Organisation for Economic Development and Cooperation.

Schultz, T.W. 1975. The value of the ability to deal with disequilibria. *Journal of Economic Literature*, 13(3): 827–846.

Seck, A. 2012. International technology diffusion and economic growth: explaining the spillover benefits to developing countries. *Structural Change and Economic Dynamics*, 23(in press).

Segerstrom, P.C., Anant, T.C.A. & Dinopolous, E. 1990. A Schumpeterian model of the product life cycle. *American Economic Review*, 80(5): 1077-91.

Sheng, Y., Gray, E.M., Mullen, J.D. & Davidson, A. 2011. Public Investment in Agricultural R&D and Extension: An Analysis of the Static and Dynamic Effects on Australian Broadacre Productivity.

Canberra, AU: Australian Bureau of Agricultural and Resource Economics and Sciences.

Simon, H.A. & Bonini, C.P. 1958. The size distribution of business firms. *American Economic Review*, 48(4): 607-617.

Solow, R. M. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3): 312-20.

Solow, R. M. 1960. Investment and technical progress. In K.J.Arrow, S.Karlin & P.Suppes, eds. *Mathematical Methods in the Social Sciences*, pp. 89-104. Palo Alto, CA, USA, Stanford University Press.

Solow, R. M. 1962. Substitution and fixed proportions in the theory of capital. *Review of Economic Studies*, 29(2): 207-218.

Solow, R.M. 2001. After 'Technical progress and the aggregate production function'. In C.R. Hulten, E.R. Dean & M.J. Harper, eds. *New Developments in Productivity Analysis*, pp. 173-178. Chicago, IL, USA, University of Chicago Press.

Spence, M. 1976. Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies*, 43(2): 217-35.

Sutton, J. 1997. Gibrat's legacy. *Journal of Economic Literature*, 35(1): 40-59.

Syverson, C. 2004. Market structure and productivity: a concrete example. *Journal of Political Economy*, 112(6): 1181-1222.

Syverson, C. 2011. What determines productivity? *Journal of Economic Literature*, 49(2): 326-65.

Terleckyj, N. 1974. *Effects of R&D on the Productivity Growth of Industries: An Exploratory Study*. Washington, DC, National Planning Association.

Traxler, G. & Byerlee, D. 2001. Linking technical change to research effort: An examination of aggregation and spillover effects. *Agricultural Economics*, 24: 235-246.

Trefler, D. 2004. The long and short of the Canada-U.S. free trade agreement. *American Economic Review*, 94(4): 870-895.

Trofimenko, N. 2008. Learning by exporting: does it matter where one learns? evidence from Colombian manufacturing firms. *Economic Development and Cultural Change*, 56(4): 871-894.

Tybout, J. R. 1996. Heterogeneity and productivity growth: assessing the evidence. In J. Roberts & J. R. Tybout, eds. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*, pp. 43-72. New York, NY, USA, Oxford University Press.

Tybout, J.R. 2000. Manufacturing firms in developing countries: how well do they do, and why? *Journal of Economic Literature*, 38(1): 11-44.

Van Beveren, I. 2012. Total factor productivity estimation: a practical review. *Journal of Economic Surveys*, 26(1): 98-128.

Van Biesebroeck, J. 2005. Exporting raises productivity in Sub-Saharan African manufacturing firms. *Journal of International Economics*, 67 (2): 373-391.

Verhoogen, E. 2008. Trade, quality upgrading and wage inequality in the Mexican manufacturing sector. *Quarterly Journal of Economics*, 123(2): 489-530.

Wallace, J. 2007. Exports and productivity: a survey of the evidence from firm-level data. *The World Economy*, 30(1): 60-82.

Wößmann, L. 2003. Specifying human capital. *Journal of Economic Surveys*, 17(1): 239-270.

Yang, Y. & Mallick, S. 2010. Export premium, self-selection and learning-by-exporting: evidence for Chinese matched firms. *The World Economy*, 33(10): 1218-1240.