Discovering the Gains from Trade: Alertness and the Extent of the Market

Adam Martin*1

1. Introduction: Ricardian and Smithian Gains

The phrase “gains from trade” perches on the lips of economists as easily as “compared to what?” or “incentives matter.” Virtually all students of economics have the phrase drilled into their heads from their first principles class. Positive sum exchange plays a central role through the whole wide world of economics, from research to pedagogy to policy debates (Caplan 2002, Whaples 2009). That trade allows for not only subjective gains in utility but also gains in the sheer quantity of goods available is one of the most trumpeted insights of economics. Among economists, it is nearly universally held not only as a theoretical possibility but as a robust concept with extremely wide-ranging applicability. The professional burden of proof falls squarely on the shoulders of those who would posit an exception, whether theoretical or practical.

The gains from trade are, arguably, the central lesson of economics as a discipline since Adam Smith explicated its causal role in determining the wealth of nations (1776) and David Ricardo (1817) detailed the principle of compara-

---

* Adam Martin is a post-doctoral fellow with the Development Research Institute at New York University.

1 Richard Wagner, Nigel Ashford, Daniel D’Amico, Nicholas Snow, and the participants in the Institute for Humane Studies Summer Research Fellowship provided helpful comments on early drafts that became this paper. I would also like to thanks Emily Chamlee-Wright and Joshua Hall for inviting me to submit the paper and present it at Beloit College. A version was also presented at Loyola University in New Orleans, where I received helpful feedback from the attendees.
tive advantage. But there exist substantial differences between the Smithian story about the extent of the market and the Ricardian tale of comparative advantage (c.f. Warsh 2006). In the Ricardian approach, natural productivity differences furnish an instrumental reason for trade. In the Smithian approach, natural trade is what generates potential differences in productivity.

Both the Smithian and Ricardian approaches have their own strengths and appeal. Ricardian thinking has traditionally dominated economists’ imaginations, probably owing to its neat conformity to the canons of rational choice. Though formulated in a pre-marginalist framework, the principle of comparative advantage carries over to modern economics nicely due to its implicit reliance on opportunity cost and emphasis on the instrumental exploitation of underlying technical factors. Since productivity differences are treated as given, such gains are dissipated by the act of specialization. In the limit, such analysis explains why agents are specialized to the extent they are rather than why they would specialize any further.

Not so with the Smithian vision. Rather than producing a static, one-time gain, trade leads to a virtuous cycle of economic development. Different relative productivities emerge from the division of labor rather than causing it. Smith argues that existing specialization prompts the discovery of better ways of doing things. He offers the example of a boy working on a steam engine who, having been assigned a particular task, figures out a simple labor saving way to automate it (Smith 1776: 20). Productivity increases emerge from the extant division of labor. The division of labor, in turn, is limited by the extent of the market. Fortunately for the inhabitants of a Smithian world, the extent of the market is not limited by an instrumental assessment of extant differences in productivity, but by the “natural propensity to truck, barter, and exchange” (Smith 1776: 25). The desire to trade is a sort of unmoved mover in Smith. Rather than dissipating, the gains from trade snowball.

To some extent, these simplified accounts are caricatures of Ricardo’s and (especially) Smith’s rich ideas. But there remains a fundamental analytic distinction between Smithian and Ricardian gains from trade. Ricardian gains are realized through the exploitation of existing productivity differences. Smithian gains are

---

2 “Given” in this context refers primarily to the standard method for teaching comparative advantage as a short run phenomenon, in which productive capacity is static. However, this ultimately also applies to approaches that treat changes in productivity as following a functional form in the long run.
the result of changing productivity. Ricardian gains are far more ably handled by standard economic theory. They are also economically significant. Cline (2004) estimates as-yet unrealized Ricardian gains owing to international trade restrictions, finding them large enough to potentially lift half a billion people above the World Bank's $2/day global poverty line. Even in the U.S., a rich nation with relatively low trade barriers, Bradford et. al. (2006) survey extant estimates and report that eliminating existing protectionist policies would result in a gain of between $400 billion and $1.3 trillion per year, which comes to $4,000-$12,000 per household.

These are hefty figures, but viewed historically Smithian gains have been much larger. McCloskey (2010) provides the most extensive survey of the evidence from economic history that increases in income per capita over the past two centuries have dwarfed plausible Ricardian gains by several orders of magnitude at least. The explosion in material well being since the Industrial Revolution has been largely the result of changing productivity, not the exploitation of given productivity differences. The economic growth literature has reflected this finding since the 1950's, even if it is only sometimes connected to Smithian gains from trade.

Modern economic theory is much more comfortable with Ricardian gains, and they are empirically quite real. But Smithian gains dwarf them. This essay investigates how one might account for Smithian gains without sacrificing the intuitive and empirical appeal of Ricardian instrumental gains. How do Smithian gains fit into a modern, instrumental logic of choice? How do the gains from trade snowball in the absence of a non-instrumental propensity to truck, barter, and exchange?

Section 2 lays out the most common extant explanation of Smithian gains with instrumental action, the Increasing Returns Approach. This approach is rooted in the literature on endogenous growth theory and turns on the non-rivalry of knowledge.

Section 3 uses Israel Kirzner’s theory of entrepreneurship to develop an alternative to Increasing Returns, which I dub the Discovery Approach. In Kirzner’s theory, entrepreneurs discover profit opportunities to which they are alert. I argue that discovery is a plausible source of Smithian-sized gains that does not abrogate the centrality of instrumental action, and further that the specialization of alertness allows those gains to generate further gains rather than dissipating.

Having laid out the broad contours of the Discovery Approach, Section 4 draws out an important implication of the theory: the Hayekian division of
knowledge is a key driver of economic development. Examining this implication in detail allows for a preliminary comparison of the relative plausibility of the Discovery and Increasing Returns Approaches. Some basic stylized facts of development, I argue, favor the Discovery Approach. Section 5 concludes.

2. The Increasing Returns Approach

In the modern literature on economic growth, all roads lead back to Robert Solow (1956). He set out to test whether capital accumulation could explain the growth in the U.S. economy by relating growth in Gross Domestic Product (GDP) to returns on capital. Land had long ago ceased to be a significant fraction of GDP, and the addition of extra labor to GDP tells us nothing about GDP per capita. What he found was that none of the returns to accumulating factors of production accounted for the overall increase in output. The bulk of growth was in the residual of his equation, the return to some omitted factor. The ordinary reshuffling and accumulation of capital, as economists have known since the marginal revolution, follows the same principles that govern Ricardian gains from trading fixed stocks. Rents, or returns to capital, are dissipated. Empirically, Solow’s findings began the process of establishing that empirically observed growth rates dwarf plausible Ricardian gains.

Momentously, Solow’s residual came to be interpreted as technological innovation (c.f. Easterly 2001: ch. 3). Of course, there is no variable to cleanly measure technological change. A cynic might believe that this is why the idea has found such unshakeable purchase. But it does make intuitive sense. If more tools were insufficient to account for increased growth, perhaps better tools could explain it. Technological change in this context means having tools made with better recipes or blueprints. It is difficult to deny that we have more effective tools than our ancestors, and increasingly so. Technological innovation is a wholly plausible explanation for Solow’s residual. This explanation also plays right into the Smithian vs. Ricardian distinction. Ricardian gains accrue to reshuffling factors of production into higher valued uses—including accumulating them—while Smithian gains are the result of increased productivity brought about by technological change.

---

3 In this section the literature I identify as constituting the Increasing Returns Approach is detailed best by Romer (1994) and Warsh (2006). For the sake of space I ignore other strands of the tradition, notably Buchanan and Yoon (1994).
Endogenous growth theory picks up where the Solow model leaves off. It can be seen as the concerted attempt by a number of economists to eliminate Solow’s residual. The residual provokes such hostility for two reasons. Despite the common acceptance that it represents technological innovation, the residual in Solow’s equation is still a black box. The model depicts technology as wholly exogenous to the economy, a thoroughly implausible characterization. Endogenous growth theory—a strange moniker given that any theory of growth would purport to endogenize growth itself—is so named because it models technological change. The dominant strain of this theory treats technology as resulting from investment, the output of a function whose inputs include human capital and previous technology.

The second impetus for excising the residual comes from an empirical puzzle it generated, convergence. The application of the Solow model to cross-country analyses of economic growth—admittedly not the end it was designed for—generated a stark prediction that countries’ growth rates would converge to a common number, namely the residual rate of GDP growth due to technology. The reason followed strict Ricardian logic: poorer nations would have lower capital stocks and thus a higher marginal return on capital. Capital should flow to poorer nations until the rates of return equalize, leaving on the residual change in technology. But technology is easy to copy since technology is just recipes for tools. Poorer countries today, for instance, can skip wired telephone lines and go straight to cell phone towers. The problem is that convergence did not happen (c.f. Easterly 2001: ch. 3). Growth rates vary a fair amount from one nation to another, and poor countries are not systematically catching up. Endogenous growth theory, then, tries to take more seriously the question of differences in cross-country growth.

The major puzzle confronting endogenous growth theory from the outset is: if technological recipes are produced by investment in research and development, how is technological change any different than capital accumulation? If technological innovation is the outcome of investment, why are the gains from innovation not dissipated just as the Ricardian gains to capital accumulation are? Endogenous growth theory attempts to escape this conundrum by ascribing to the production of technological knowledge what Kenneth Arrow dubbed the “under-

---

4 Olson (1996) points to some more systematic empirical support for the idea that technological knowledge is extremely cheap to import and thus not a constraint on growth.
ground river” in economics, the idea of increasing returns (Arrow 1994). Increasing returns to technological knowledge—new recipes, blueprints, etc.—means that the more technology you have, the cheaper it is to develop new technology. Since existing technology is an input into future technology, the rate of technological innovation should increase over time leading to exponential increases in GDP. Since this is the primary mechanism by which endogenous growth theory attempts to introduce Smithian gains into contemporary Ricardian economics, I refer to it as the Increasing Returns Approach.

The idea of increasing returns predates endogenous growth theory, but economists were typically hesitant to employ it until about three decades ago (c.f. Young 1928). The reason was that if output became cheaper to produce the more one already produces, then one firm with an early start could easily monopolize an industry by undercutting its competitors (c.f. Warsh 2006: 46). But where are all these monopolies, especially in high tech industries? To solve this conundrum, the Increasing Returns Approach leans on the non-rivalry of knowledge. If you utilize a piece of knowledge, there is no less of that knowledge for me to utilize. This is especially important in the formation of new ideas. Ideas are inputs into future ideas. But an idea that you have can serve as an input into my future ideas as well as your own with no tradeoff between the two.

But if knowledge is non-rival, the convergence conundrum rears its ugly head again. Utilizing technology does not require having developed the technology. That includes utilizing it for developing more technology. The Increasing Returns Approach thus looks appealing when one only looks at the historical trend of average world GDP. To the extent that it does matter empirically, that long-run global trend is its most natural and convincing application. But what about cross-country differences?

To explain why some nations are rich and others are poor, the Increasing Returns Approach relies on neighborhood effects (c.f. Warsh 2006: 80-81). Neighborhood effects, in this context, are localized knowledge spillovers. These are spillovers that only benefit those in a limited proximity. The classic example is Silicon Valley. A high concentration of computer and software engineers allows for the transmission of knowledge and problem-solving techniques to spillover from one firm to another, whether its through employees in different firms interacting with one another or employees’ ease of movement from one firm to another.

These neighborhood effects allow for an industry in a given area to innovate at a rapid pace, capturing a larger portion of the returns to that innovation.
Increasing Returns take hold when the payoff from existing technological know-how and human capital increases simply by being in proximity to others similarly endowed. Because that effect is recognized, individuals with high-tech human capital will tend to congregate in some areas, while individuals with low-tech human capital will congregate elsewhere. Since technological knowledge and human capital are inputs into innovation, those areas with concentrations of high-tech human capital will grow more quickly and become wealthier over time in an explosion of Smithian gains from increased productivity.

3. The Discovery Approach

Two signature concepts highlight the distinctive features of Kirzner’s theory of entrepreneurship: discovery and alertness. Entrepreneurs discover profit opportunities because they are alert to them. In the context of a market economy, a “discovery” takes place when an entrepreneur becomes aware of a new possibility for mutually beneficial exchange (Hayek 1968, Kirzner 1997a). That is, he recognizes a Ricardian gain waiting to be grasped. Prior to the discovery, the entrepreneur is in a state of “sheer ignorance” concerning the potential trade (Kirzner 1997a). Alternatively, the discovery can be labeled “sheer creativity,” but it makes little difference *qua* economic theory whether opportunities are “out there” waiting to be discovered or are genuine ontological novelties. What matters is that some entrepreneur is in a position to identify hitherto unrealized gains. And while many potential gains will turn out to be illusory, for the purposes of what follows (and for reasons that will be made clear) a discovery is taken to mean a genuine profit opportunity that pans out for the entrepreneur who follows through on it.6

Kirzner often appends the adjective “costless” to the notion of discovery (c.f. Kirzner 1973: 226-7). This strikes some economists as strange, given our predilection for insisting that free lunches are the most mythical of creatures. Economics is about choices between options, and options have costs. What Kirzner appreciates is that options have to be grasped by the mind before they can be chosen. Discovery is costless because it is logically prior to choice. The cost of a

---

5 This section draws extensively on Martin (2011). In that paper I deploy Frank Knight’s (1921) definition of uncertainty as a more apt description of the ignorance to which (Kriznerian) entrepreneurship is a response. That argument is broader, discussing social processes more generally. Since this article seeks to answer a narrower question these points have been summarized here.

6 Kirzner (1997b) makes clear that in speaking of discovery as perception he is using a metaphor, and is not taking a position on whether the future is “a rolled-up tapestry.”
pure discovery is not zero; it is null. This distinguishes it from Stigler’s conception of search (1961), which is a choice to look for a given piece of information. Search is costly because search is an activity that is chosen. But a discovery cannot be chosen because, by definition, one is not aware that a given option exists before it is discovered.

Distinguishing discovery from search does not imply that discovery must be wholly accidental. The two concepts are analytically distinct but may be empirically bundled. Just as a wage earner in a modern economy is earning returns on both labor and human capital, searching may prompt discovery. In fact, while discovery is not itself a chosen course of action, it always occurs while engaged in some action. Time does not stop. Individuals are always allocating their scarce time to some costly activity or another. Some activities may in fact entail a greater propensity to discovery. Admitting this may seem to blur the distinction between discovery and search, but crucial differences remain.

Even though costly activities can be intentionally undertaken as precursors to discovery, the returns to a discovery can be wholly disproportionate to the cost of the precursory activities. Pure search assumes some expected value to the information sought. But when the nature of the knowledge is unknown beforehand—that it is this or that sort of opportunity, for instance—then the expected value is unknown. The expected value of a car cannot predate the idea itself of a car. What this means is that the returns to discovery do not follow the normal principles that govern returns to factors of production (whereas the returns to search do). That is exactly what makes discovery a good candidate for explaining economic development.

McCloskey (2010) argues forcefully that factor accumulation cannot explain the remarkable growth of the past few centuries. Ricardian instrumental trade, which includes allocating factors to production, produces growth several orders of magnitude too small to explain the modern world. The payoffs, thanks to competitive market pressures, are always proportionate to the cost. Discovery severs the link between cost and reward, creating space for much larger increases in the

---

7 Evans and Friedman (2011) make a helpful distinction between searching, in which one knows what one is looking for, and browsing, in which one is looking but not for a particular piece of information. Holcombe (2007: 64-5) argues that, while technological innovations are Kirznerian discoveries, research and development activities increase the propensity to make such discoveries.

8 Kirzner (1973: 66) argues that entrepreneurship must be distinguished “sharply” from factors of production.
size of the overall economic pie. McCloskey argues that something like discovery must be taking place in order to explain growth beyond factor accumulation. Discovery—like the increasing returns analogy—drives a wedge between instrumental action and rent dissipation. Already discovered profit opportunities are competed away after their initial implementation, but there remains the possibility of new discoveries.

I thus refer to the attempt to (however partially) explain how the gains from trade relate to economic development with Kirzner’s theory as the Discovery Approach.9 Fundamentally, the Discovery Approach differs from the Increasing Returns Approach in how it explains super-Ricardian gains from trade. Ricardian trade, recall, is instrumental choice. Increasing Returns identifies post-choice spillovers as the source of explosive gains. The Discovery Approach instead identifies pre-choice creativity or perception as the source. The gains from discovery cannot be normalized by competition until after the discovery is implemented.

To move from individual discoveries to a full-blown Smithian process, however, a bit more is needed. One must explain why discoveries lead to further discoveries. Without such chain reactions, discovery just becomes another source of exogenous shocks.

Martin (2011) offers a taxonomy of different forms that discovered gains from trade might take. I expand that analysis here, classifying entrepreneurial gains from trade as Adjustments, Producer Innovations, and Consumer Innovations. These forms are analytic; like discovery and search, they may be empirically bundled. By definition, they all involve a pure economic profit and thus occur before other entrepreneurs have arbitraged away the profit opportunity.

Adjustment is when the entrepreneur recognizes or anticipates a shift in consumer demand or a supply shock. He reallocates resources from a lower to a higher valued use, but that is more or less the end of the story. Pure Adjustment involves no new goods or services. Other factors of production may have to be re-shuffled on account of changing prices. But beyond the initial discovery, Adjustment is just Ricardian rationality responding to a change in underlying variables. New discoveries may happen, but under an assumption of consumer sovereignty

---

9 Kirzner himself uses the phrase “discovery approach” to refer to his particular concerns in the Austrian theory of the market process (Kirzner 1997a). My use is far more narrow and idiosyncratic.
there is no obvious mechanism to spur them on; they would be exogenous vis-à-vis the initial discovery.10

Producer Innovations are discoveries of lower cost methods of bringing existing kinds of consumer goods to market. Those methods might involve more efficient technologies, but could also involve the discovery of new trading opportunities, new resources, or new organizational methods. “New” in this context includes both new local applications of non-rival ideas and globally new ideas. Such discoveries free up resources for use in other lines of production. They also increase real consumer income with which to purchase the fruits of that production. Releasing resources from other lines of production decreases their price by eliminating their previously most-valued use, lowering the cost of pursuing other opportunities. But while Producer Innovation allows other entrepreneurs to pursue projects they have “waiting in the wings”—discovered but not yet acted upon—it does not in itself give a reason to suspect more discoveries. After the initial discovery we are back to Adjustment, wherein other entrepreneurs respond to the innovation as a supply shock. Adjustment, as we have seen, is not self-reinforcing.

Consumer Innovations are discoveries of new consumer (or first order) goods. As with a Producer Innovation, a Consumer Innovation only needs to be locally new in order to turn a profit. Consumers must value the new good more than some other good that they spend their income on and more than the alternative uses of the inputs used to create it. But it need not free up resources on net; a greater benefit may well be worth more overall resource expenditure. Such a discovery would constitute psychic but not material gains: not more, but more highly valued goods. Displacement of close market substitutes may free up the inputs used to produce them, but producing a new good requires that inputs be diverted from other lines of production. The net effect on resource consumption could go either way. In either case, the process leads back to Adjustment and fails to provide a reason to expect more (or fewer) discoveries. None of the three forms that a discovery might take, then, seems capable in itself of accounting for further discoveries. Kirznerian alertness, I posit, can.

10 It is possible that shifts in consumer demand lead to further shifts in consumer demand, such as with network goods, but this is making an idiosyncratic technical assumption the driver of change. While this sort of change is a possibility, the purpose of this essay is to find a more robust explanation that can explain more widespread patterns of development across different types of goods.
Alertness is the second signature concept characterizing Kirzner’s theory of entrepreneurship (Kirzner 1973: 35ff, 65ff). It is “the human propensity to sniff out opportunities lurking around the corner” (Kirzner 1979: 29). That is, it is the propensity to make discoveries. Discovery, Kirzner surmises, is not wholly a matter of luck (Kirzner 1997a: 72). It does involve being in the right place at the right time, but also depends on the entrepreneur’s ability to correctly perceive the profit opportunity. Alertness might involve pattern recognition, data interpretation, signal extraction, imagination, or some other cognitive function that does not reduce to merely gathering information. What matters is that discovery is neither automatic nor deliberate. One can wish to recognize profit opportunities but cannot do so by sheer act of will. Discovery with any reliability requires honed alertness.

Kirzner’s “propensity” to discover serves a function similar to Smith’s “propensity to truck, barter, and exchange.” Both propensities serve as engines driving a process forward in time. Alertness drives the market process toward mutual plan coordination; without it, discovery might be no more likely than error, belying the order that we in fact observe in the marketplace. Smith’s propensity to trade drives the extension of the market, allowing for the further division of labor. The crucial difference is that alertness is an epistemic engine rather than a motivational one. Knowledge drives the process rather than incentives. This is why it can bridge the gap between Ricardian gains and Smithian gains: Ricardian instrumental rationality need not be overridden in order to endogenously generate ongoing activity and change.

Alertness is a sort of unmoved mover in Kirzner’s major works. He takes it as a given (Kirzner 1992: 26) because the puzzle he seeks to explain is the proximate cause of plan coordination in the market process. While this works well for disentangling the puzzle of coordination, we would go wrong in applying Kirzner’s framework to other puzzles if we were to mistake “given” for “constant.” Alertness is exogenous in Kirzner’s work; it need not be fixed when extending it. Endogenizing alertness allows Kirznerian entrepreneurship to account for long-run economic change as well as short-run adjustment, bridging the gap between the Ricardian and Smithian gains from trade.

One key feature of alertness is that it is not a general ability to perceive profit opportunities. All agents “share in this ability to some extent… [b]ut some have higher degrees of this ability—some in some lines of endeavor, others in other lines of endeavor” (Kirzner ibid.). Alertness is specialized to a context of action.
If we make the plausible assumption that experience in a “line of endeavor” specializes individuals’ alertness to the context of that endeavor, then we can explain how discoveries can lead to further discoveries. A discovery can lead to further discoveries when it allows an entrepreneur’s alertness to be further specialized. Hayek may have had this sort of cognitive specialization in mind when he wrote:

> Competition is as much a method for breeding certain types of mind as anything else: the very cast of thinking of the great entrepreneurs would not exist but for the environment in which they develop their gifts. The same innate capacity to think will take a wholly different turn according to the task it is set (Hayek 1979: 76).

Of the three forms that discovery might take, Producer Innovation most clearly opens the door to ongoing gains from trade. A cost-saving discovery frees resources for entrepreneurs to pursue new lines of endeavor; it is only in actual pursuit that alertness is cultivated. Imagination and forecasting are a function of present alertness. Until the context of a new production process or new market activity is lived in, it cannot be an incubator of new alertness. That requires an actual investment of resources and especially time. By freeing resources for use in other entrepreneurial plans, Producer Innovation enables the cultivation of alertness that might not have come to be otherwise.

Adjustment and Consumer Innovation can likewise open up new contexts that are alertness incubators. In their pure form, however, they do this by shuffling resources out of other activities. Specialized knowledge can lose the context of its usefulness. Producer Innovation is different because, by providing some service or a close substitute at lower cost than before, it allows for increased resources to be channeled into other lines of endeavor without giving up production of the service thus innovated. That is, Producer Innovation allows for more intensive specialization and systematic increases in the economy’s diversity of enterprises. Discoveries of lower cost means of production enable increases in the extent of the market. Increasing the extent of the market increases the scope for specialized alertness to develop, prompting more discoveries. We are back to a Smithian understanding of specialization and the gains from trade, which is depicted in Figure 1:

---

In genuine market competition, of course, this is only because such specialized alertness noticed less valuable profit opportunities than what displaced it.
Alertness and the Extent of the Market

Entrepreneurial Discovery

Resource Reallocation

Altered Alertness

Figure 1: The Discovery Approach

The Discovery Approach goes as follows: an entrepreneur discovers a lower cost means of production and reallocates resources accordingly, earning an entrepreneurial profit. Doing so frees up resources for entrepreneurs to initiate projects that were previously too expensive. In so doing, the alertness of those entrepreneurs is altered, making them more adept at discovering certain types of profit opportunities. Some of those discoveries may themselves involve lower cost means of production, allowing the whole cycle to repeat. In Smithian fashion, existing specialization prompts the discovery of new exchange opportunities. In Ricardian fashion, that existing specialization comes about not through some natural propensity, but through the instrumental allocation of resources to exploit a comparative advantage. The gains from trade are twofold. The Ricardian gains from trade occasion Smithian gains by allocating freed resources to a new context and inculcating alertness.

The basic features of the Discovery Approach are not new. Technical elements of Kirzner’s theory are deployed in new ways to explain the specific puzzle of development, such as discovery being a source of explosive returns and alertness being the relevant sort of specialized knowledge. Smith’s child innovator notices the profit opportunity in automating the steam valve because of the task to which his mind is set. That trade is both a cause and consequence of the division of labor was recognized at least as far back as Mises in Socialism:

Had the strength and abilities of all individuals and the external conditions of production been everywhere equal the idea of division of labor could never have arisen. Man would never of himself have hit upon the
idea of making the struggle for existence easier by co-operation in the division of labor…

Once labour has been divided, the division itself exercises a differentiating influence. The fact that labour is divided makes possible further cultivation of individual talent and thus co-operation becomes more and more productive. Through co-operation men are able to achieve what would have been beyond them as individuals, and even the work which individuals are capable of doing alone is made more productive (Mises 1922: 260).

4. The Division of Knowledge and the Wealth of Nations

Civilization advances by extending the number of important operations which we can perform without thinking about them. Operations of thought are like cavalry charges in a battle—they are strictly limited in number, they require fresh horses, and must only be made at decisive moments (Whitehead in Hayek 1960: 22).

Hayek described the “division of knowledge” as “quite analogous to, and at least as important as, the problem of the division of labor” (Hayek 1937: 39). The knowledge embedded in social processes continually expands in developing economies, but the percentage of that knowledge that an individual can command grows ever smaller. Individuals’ knowledge of productive activities is ever-increasingly differentiated and specialized. Leonard Read (1958) famously described how no one mind could master all that was necessary to make a pencil. The division of knowledge is central to the Discovery Approach. Local knowledge of time, place, and particular circumstances enables the development of more finely honed alertness and thus an increased propensity to discover gains from trade in the context of that locality.

The Discovery Approach’s emphasis on the differentiation and division of knowledge stands in contrast to the Increasing Returns emphasis on technology, defined as non-rival and thus replicable knowledge. This of course grew out of the distinction with scarce capital from the older Solow model. Increasing Returns, like the Solow model before it, relies on the possibility of re-using good ideas to escape from the Ricardian box. Technological change is just changing recipes for
combining existing materials, and recipes can be copied. That is, ideas cause progress to the extent that they can be held in more than one mind.

While the non-scarcity of ideas is beyond doubt, it is a strange thing to identify as the source of growth when no one mind contains the knowledge necessary to create a pencil (Wagner 2006). This is not to deny that technology is, empirically, a vital part of development. But this is a long way from saying that non-rivalry of knowledge is the crucial component of Smithian gains. The Discovery Approach diverges in several important ways from such a conclusion:

- Even when technological recipes spread, they must be adapted to local conditions and particular production processes (Holcombe 2007: 80).
- Specialization is neither a one time phenomenon nor a continuously separating equilibrium between some predefined set of goods and services, but rather an ongoing process that involves structural, morphological change to the economy’s capital structure and increasing social complexity (Lachmann 1956, Beinhocker 2006).
- The salient relationship between knowledge and development is not aptly characterized by an ever-wiser representative agent or firm but by the growing relative ignorance and increasingly heterogeneous knowledge of all members of society (Hayek 1960).

For the Discovery Approach, the crucial fact is not the non-rival possession of knowledge but the non-rival use of knowledge. Put differently, it is not because knowledge can be possessed by multiple individuals but that it need not be in order to benefit all of them that is key to development. Through trade, individuals make use of others’ knowledge in the pursuit of their own ends. This frees them to engage in more specialized tasks and thus accumulate more refined knowledge of the tasks in which they engage.

That ever increasingly refined knowledge can be embodied technologically, but it need not be. From the perspective of allowing others to further specialize, the usefulness of the replicable knowledge embodied in a frozen dinner on the one hand and of the idiosyncratic, local, and tacit knowledge of a cook using his own specialized tools on the other is no different in principle. Benefitting from others’ knowledge is the general form of prosperity; technology is but one form that such prosperity can take.

Another important aspect of divided knowledge is heterogeneity. Non-rivalry of knowledge only drives development to the extent that it is useful for
many individuals to have the same kind of knowledge. The Increasing Returns Approach is predicated on the idea that strength builds on strength. Endogenous growth theorists such as Paul Romer latch onto examples like Silicon Valley and Hollywood. Jane Jacobs advances an understanding of urban growth more focused on the division of knowledge. For Jacobs, it is because different sorts of work exist side by side in cities that they are engines of growth. Glaeser et. al. (1992) test these implications, finding that Jacobs was right. Cities with more diverse industries grow faster than those with more concentrated industries. Today’s Silicon Valley can be tomorrow’s Detroit.

Moving beyond cities, the emphasis on divided knowledge over non-rival knowledge also leads the Discovery Approach to locate Smithian vs. Ricardian gains at different scales of interaction than does the Increasing Returns Approach. Increasing Returns treats Smithian gains as local and Ricardian gains as global. Recall that endogenous growth theory identifies Smithian gains with neighborhood effects in order to explain cross-country income differences. Larger gains are due to localized knowledge spillovers, while smaller Ricardian gains obtain between trading areas (e.g., nations). More precisely, they obtain between representative agents of the two areas. The Discovery Approach, by contrast, treats Ricardian gains as local and Smithian gains as global. Ricardian gains obtain between the two parties to a trade. Smithian gains, by contrast, are a property of the global network of trading relationships that maps onto the social division of knowledge. Ricardian gains are realized in individual markets, Smithian gains in the market process as a whole.

Olson (1996) points to some stylized facts that favor the Discovery Approach over Increasing Returns on this margin. He notes that, according to endogenous growth theory, the wealthiest nations should also grow the fastest. By definition, increasing returns mean that the more knowledge one has the easier it is to accumulate yet more. But while rich countries have better growth rates on average than poor countries, they are not the fastest growing. The fastest growing countries at any given time are a subset of poorer nations. Olson speculates that these rapid growth areas are recently liberalized economies eliminating (Ricardian) inefficiencies, a speculation consistent with evidence on international trade liberalization as well as broader measures of economic freedom.

Olson’s insight is consistent with the Discovery Approach’s identification of Ricardian gains as local and Smithian gains as global. In the last few decades, the Smithian gains from trade emerge from an international division of knowledge.
Tapping into that global network of dispersed knowledge allows a country to grow faster than if it was left to develop on its own. The initial gains will be huge as one establishes first (or more intensive) contact with an existing division of knowledge. But accessing that network requires trading with those who already participate in it, in the same way that the existing stock of information on the internet requires a computing device with a network connection. Barriers to realizing Ricardian gains can thus limit the realization of Smithian gains as well, an implication that fits Olson’s observations on growth rates precisely. The Smithian story tells us why such incredible material gains have been realized over the past few centuries; the Ricardian story tells us why those gains do not spread to all corners of the globe.

5. Conclusion: Economic Development as an Economic Problem

As Adam Smith famously observed, the division of labor is limited by the extent of the market. It should be abundantly clear that, in spite of borrowing Smith’s rhetoric, the Discovery Approach fundamentally differs from Stigler’s (1951) classic re-interpretation of the two halves of that statement. “The division of labor” or “knowledge” is not the division of a given task, but the division of the multitude of varied and often incompatible plans. Similarly, “the extent of the market” does not refer to the demand for a particular product but to the scale and scope of the overall nexus of trade. It is because some have specialized in a given task that others can specialize in different tasks. Ricardian gains explain when individuals will find it in their interest to further specialize. Smithian gains are the product of the resultant specialized alertness.

And yet, Stigler’s appropriation of Smith’s dictum, in contrast to the Discovery Approach, analogizes Smith’s insight to a production function for a single firm. Stigler’s interpretation is the dominant one.12 While Smith did use the pin factory as an example of the effects of dividing labor, his discussion is far broader. The passage about the boy and the steam engine moves from common laborers to machinists to “philosophers and men of speculation” (Smith 1776: 21). This is Hayek’s extended order, not a factory floor. In pointing out this gap, I am not so

---

12 Warsh, for instance, tries to pigeonhole Smith into a “contradiction” between the invisible hand of perfect competition and the pin factory’s increasing returns (Warsh 2006, p. 47). Neither of these is a good characterization. The invisible hand is better described by Kirzner’s market process than by perfect competition (c.f. Holcombe 2007, pp. 51-4), and the pin factory is not the sum total of the division of labor.
much attacking Stigler’s argument as much as theories that make a similar move with respect to the entire economy. Economic development is poorly described by a production function. Solow’s original findings should have freed economists from that dementia, but instead they fashioned it into a straight jacket. Development economists’ solution to Solow’s finding that firm-like factor accumulation could not explain growth patterns was to imagine the economy as an industrial park of several similar firms and tweak the production function accordingly. From the perspective of the Discovery Approach, this solution is far from satisfying.

The fundamental problem with the Increasing Returns Approach is that it analytically reduces a Smithian process to a Ricardian allocation. The divergence between these two stories about the gains from trade connects to a deep and wide gulf between two analytic visions in economics (c.f. Wagner 2004). The Ricardian vision is one of allocative response to given constraints. The Smithian vision is of an open-ended process of ongoing social interaction and individual transformation. Both of the visions can shed light on aspects of the social world, but to force a Smithian phenomenon into a Ricardian box is to rob it of the features that give it unique explanatory power. The dominant paradigm in economics is to reduce all facets of social reality to a Ricardian allocation. This is a mistake. The two visions can be connected—this is what the Discovery Approach tries to do—but one cannot be made to do the work of the other.

In addition to its analytic niceties and pedagogical advantages, the Ricardian vision has two viscerally appealing features. First, it allows the economist to tell stories with a beginning, middle, and end. In the beginning was the productivity difference, followed by specialization and production, all of which is consummated in trade until the zero profit condition obtains. Second, it explains why it makes sense to trade from an individual, instrumental point of view. We trade because we expect it to make us better off. Importantly, the Ricardian approach also tells us why we stop trading: when we think that is no longer the case.

The appeal of the Smithian vision rests in its open-endedness. That open-endedness emerges from its more strictly social character, by which I mean that the Smithian emphasis on interaction (rather than just action, such as allocation) highlights the role of inter-individual forces. A Smithian approach, unlike a Ricardian one, would never see a fundamental equivalence between Robinson Crusoe and an entire economy. Representative agents are the opposite of Smithian individuals, who are differentiated by various social processes including the division of labor. When economists lose sight of the interactive element of exchange, they
lose sight of the distinction between technical problems and economic problems (Buchanan 1964).

Economic development is an economic problem, one that can only be “solved” by a genuinely Smithian process. Robust long-run development involves coordinating a continuously differentiating division of knowledge, an insight that should be obvious when economists cease modeling economies as production functions. But the important insights of the Ricardian vision—those that arise from its application to individual behavior, rather than its misapplication to entire economies—should not be discarded. Israel Kirzner’s signature concepts of entrepreneurial alertness and discovery allow us to break out of the Ricardian box and into a larger Smithian world without abandoning those Ricardian insights. In this as in many other economic and social puzzles, the usefulness of Kirzner’s ideas reaches far beyond the tasks to which he put them. Economists wishing to discover new insights would do well to hone their alertness to his work.

References


_____. 1997b. Interview with Israel M. Kirzner. The Austrian Economics Newsletter 17 (1).


