



Econ Journal Watch
Scholarly Comments on Academic Economics
Volume 1, Issue 3, December 2004

COMMENTS

Adam Smith's Invisible Hands, *Peter Minowitz* [381-412](#)

'Credibility' in Context: Do Central Bankers and Economists Interpret the
Term Differently? *James Forder* [413-426](#)

Identity and Politics in School Reform Research
Fabio Rojas [427-436](#)

Fashion Cycles in Economics
Philip R. P. Coelho, Daniel B. Klein, and James E. McClure [437-454](#)

Response to Coelho et al., *Wolfgang Pesendorfer* [455-464](#)

Rejoinder to Edwards and McGuirk, *Jib Y. Chang and Rati Ram* [465-472](#)

CHARACTER ISSUES

Scholasticism vs Pietism: The Battle for the Soul of Economics,
Robert H. Nelson [473-497](#)

INVESTIGATING THE APPARATUS

The National Research Council Ranking of Research Universities: Its
Impact on Research in Economics, *Randall G. Holcombe* [498-514](#)

ECONOMICS IN PRACTICE

Economics without Entrepreneurship or Institutions: A Vocabulary
Analysis of Graduate Textbooks, *Dan Johansson* [515-538](#)

CORRESPONDENCE [539-545](#)

Adam Smith's Invisible Hands

PETER MINOWITZ*

A COMMENT ON: WILLIAM D. GRAMPP. 2000. WHAT DID ADAM SMITH MEAN BY THE INVISIBLE HAND? *JOURNAL OF POLITICAL ECONOMY* 108 (3): 441-465.

[Abstract, Keywords, JEL Codes](#)

“WHAT DID ADAM SMITH MEAN BY THE INVISIBLE HAND?”

William D. Grampp poses this long-disputed question and answers it presumptuously via his article in the *Journal of Political Economy*. In trying to constrain the reach of Adam Smith's invisible hand, Grampp offers this summary of what it is, and what it is not.

True, the invisible hand does have a consequence that is unintended, but the consequence is not a beneficial social order. It is a benefit that, while important, is of a lesser order. It is to contribute to the defense of the nation. It is nothing so complex and so grand as the social order or the price mechanism within it. (Grampp 446)

Grampp merits approbation for his sensitivity to sometimes-neglected puzzles in Smith and for warning against the common tendency to “see” an invisible hand any time Smith argues against governmental regulation. Grampp imaginatively confronts some widely held views, wisely reminds us of Smith's departures from *laissez-faire*, and courageously accuses

* Political Science Department, Santa Clara University.

I would like to thank Murray Dry, Joan Robins, and William Sundstrom for their helpful comments.

Smith of forgetfulness, inconsistency, implausibility, irrelevance, and other shortcomings.

Unfortunately, Grampp also conveys oversimplifications, exaggerations, and distortions that represent a long backward step in Smith studies. Grampp attempts to trivialize the invisible hand and to belittle the competence of its creator. By publishing this article at the dawn of the new millennium, the *Journal of Political Economy* suggests how far the discipline of economics may be from fathoming its origins and even its presuppositions.

To combat Grampp's iconoclastic agenda, I shall present a detailed elaboration of Smith's three references to an invisible hand. After criticizing Grampp's attempt to narrow the grasp of the invisible hand within *The Wealth of Nations* (WN), I turn to his account of the invisible hand in Smith's other book, *The Theory of Moral Sentiments* (TMS). Although Grampp's interpretation of this book errs palpably, it raises questions that can help us fathom the long-disputed tension—about the worthiness of wealth and the plight of the poor—between *The Theory of Moral Sentiments* (1759), which extols God along with love and benevolence, and *The Wealth of Nations* (1776), which expels God and emphasizes self-interest. I conclude by addressing the posthumously published essay in which Smith attributes belief in invisible hands to superstitious “savages” and thus seems to impugn the appeals to an invisible hand in his own books.

“IN MANY OTHER CASES”

However tempting it is to regard the invisible hand as a metaphor/simile for Smith's whole project, Grampp prudently focuses our attention on the precise context in which the invisible hand manifests itself. He concludes that the invisible hand does *not* have “a principal place” or even a “salient” one in *Wealth of Nations* (442).

The key chapter—“Of Restraints upon the Importation from foreign Countries of such Goods as can be produced at Home” (IV.ii)—is the first of a series in Book IV that criticize mercantilist policies. Here are the three sentences that launched the invisible hand.

As every individual . . . endeavours as much as he can both to employ his capital in the support of domestick industry, and so to direct that industry that its produce may be of

the greatest value; every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the publick interest, nor knows how much he is promoting it. By preferring the support of domestick to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. (*WN* 456)

For Grampp, the unintended public benefit the invisible hand promotes is the *domestic* build-up of capital (Grampp 452). His Abstract goes so far as to assert that Smith's invisible hand is "simply" the "inducement a merchant has to keep his capital at home, thereby increasing the domestic capital stock and enhancing military power" (441).

Earlier in the chapter, Smith laments that import restrictions create monopolies (for domestic producers) that channel a society's capital in sub-optimal ways. The typical reader of *Wealth of Nations* understands Smith's point that a capital owner, by directing his industry "in such a manner as its produce may be of the greatest value . . . intends only his own gain." Grampp is right to observe that this chapter emphasizes the owner's incentives to deploy capital domestically. Smith states that, upon "equal or nearly equal profits," any wholesale merchant "naturally prefers the home-trade to the foreign trade of consumption, and the foreign trade of consumption to the carrying trade." Smith offers several plausible reasons in explaining the merchant's posture: among other things, the merchant can more easily know "the laws of the country from which he must seek redress" and "the character and situation" of the people he has to rely upon (*WN* 454).

Grampp carefully summarizes nine ways that scholars have interpreted the invisible hand; he faults all of them for perceiving an invisible hand in other situations Smith describes whereby someone "intends only his own gain" but ends up producing benefit to others. For Grampp, by contrast, an invisible hand "guides a merchant only when circumstances induce him to keep his capital at home" (447). One prominent obstacle Grampp must confront is Smith's statement that an invisible hand operates "in many other cases" to promote an end that the

relevant agent did not intend. Grampp's response is unpersuasive, not least because it is convoluted.

Does the word "cases" mean there are transactions, other than placing capital in competitive domestic trade, that add to domestic wealth and to defense? Or does "cases" mean that transactions that place capital in domestic trade contribute to something other than defense, for example, to what he calls elsewhere the "greatness" of the nation? Or does the word have all three meanings? (Grampp 452)

Let me offer a guess about what Grampp here envisions as the *three* "meanings" that "cases" can have: capital allocated to competitive domestic trade; other "transactions" that promote domestic wealth and defense; capital, allocated to competitive domestic trade, that contributes to national greatness or another public end (beyond national defense).

I credit Grampp for emphasizing the rhetorical weight Smith puts, in the build-up to the invisible hand, on fear of capital flight, but Grampp neglects three aspects of the chapter that inspire many readers to conceive the invisible hand more broadly. First, although the paragraph emphasizes the allocation of capital—an activity that some people, e.g., "those who live by wages" (*WN* 86, 266) are not equipped to undertake—the quoted section begins with two references to "every individual," including the remarkable claim that "every individual" (not just every merchant or investor) "necessarily labours to render the annual revenue of the society as great as he can." Second, the paragraph concludes with Smith stating, "I have never known much good done by those who affected to trade for the publick good" (456). If Grampp's interpretation were correct, the paragraph should instead conclude with Smith saying, "I have never known much good done by those who affected to trade to augment domestic capital and thereby promote national defense." By here questioning the accomplishments of individuals who claimed that they were trading to promote "the publick good" generally, Smith suggests that an invisible hand may operate to produce a *variety* of public benefits.¹ The conclusion of the paragraph

¹ When Smith, via the pronoun "I," makes himself conspicuous in his paragraph on an *invisible* hand—and when he invokes what he *knows* about consequences of which the immediate actors are *ignorant*—he encourages readers to pay special attention. It remains true that the *clause* containing the invisible hand refers to "an end that was no part of his intention" without specifying that this end involves benefit to the *public*. This fact, however,

establishes a contrast between the failure of merchants who intended to promote the common interest and the success of merchants who intended to promote only their own interests. Third, a few pages earlier the chapter seems to anticipate the invisible hand with a paragraph that ignores the distinction between domestic and foreign investment.

Every individual is continually exerting himself to find out the most advantageous employment for whatever capital he can command. It is his own advantage, indeed, and not that of the society, which he has in view. But the study of his own advantage naturally, or rather necessarily leads him to prefer that employment which is most advantageous to the society. (*WN* 454)

Smith does proceed to elaborate the two prongs that Grampp stresses: that “home” is “the center, if I may say so, round which the capitals of the inhabitants of every country are continually circulating, and towards which they are always tending” (*WN* 455); and that in pursuit of profit the owner will seek to maximize the productivity of his capital.

In the paragraph that immediately follows the invisible hand, Smith provides another strongly worded claim that reinforces his commitment to economic liberty.

What is the species of domestick industry which his capital can employ, and of which the produce is likely to be of the greatest value, *every* individual, it is evident, can, in his local situation, judge *much* better than *any* statesman or lawgiver can do for him. The statesman [*sic*], who should attempt to direct private people in what manner they ought to employ their capitals, would not only load himself with a most unnecessary attention, but assume an authority which could safely be trusted, not only to no single person, but to no council or senate whatever, and which would nowhere be so dangerous as in the hands of a man who had folly

supports the common view—that the invisible hand is a pivotal concept in *WN*—rather than Grampp’s attempt to narrow the hand’s reach.

and presumption enough to fancy himself fit to exercise it.
(WN 456, emphasis added)²

Thus, even in the immediate context that Grampp emphasizes,³ Smith provides ample provocation for extending the application of the invisible hand. At several points, ironically, Grampp himself offers a ridiculously universalized statement, as if led by Smith's authorial hand to overuse the word "every" and thus exaggerate the scope of the invisible hand's benevolence. According to Grampp, Smith summons the invisible hand when describing a "condition . . . in which a man who intends to benefit only himself in a particular way may, in the act of procuring that benefit, produce a benefit of a different kind for *everyone* including himself" (Grampp 443, emphasis added).⁴ Even confining our attention to the domestic front, it is difficult to specify a commercial transaction that would yield a benefit for a nation's entire population. Smith in IV.ii does use a variety of terms in describing large groupings of people,⁵ and praise what "every" *individual* can contribute by seeking profitable investments.⁶ But the

² Smith uses similar terminology later in Book IV when he states that "the law ought always to trust people with the care of their own interest, as in their local situations they must generally be able to judge better of it than the legislator can do" (531). His main targets here, however, are laws that required farmers to sell their grain directly, without the intermediation of dealers; there's nothing about a merchant keeping his capital at home (Grampp 447) and thus promoting national defense (Grampp 441, 443). Contra Grampp, it seems natural for the reader here to recall the invisible hand that Smith earlier invoked to discourage legislators from meddling.

³ While Grampp concedes that many common reflections about the invisible hand are "related to" ideas that are "in the *Wealth of Nations*, somewhere or other," he complains that these ideas typically are not "ideas that Smith himself made a part of it" (Grampp 442). I'm not sure how one can definitively specify the ideas that Smith "made a part of" the invisible hand, but one should at least scrutinize the chapter in which the invisible hand appears. When Grampp turns to *The Theory of Moral Sentiments*, as I'll shortly elaborate, his reading of its "invisible hand" passage is embarrassingly lazy, in part because he ignores the profound questions Smith poses nearby.

⁴ Grampp similarly misuses the term "everyone" on pp. 450, 451, and perhaps 459, though he provides a subtler overview on p. 444.

⁵ In WN IV.ii, Smith refers to the proper names of several European nations and peoples; he also employs the following "collective" terms: kingdom, society, country, state, "the interest of a nation," "the publick interest," "the publick good," "the circumstances of the people," "the general good," "our manufacturers," and "us." The last two phrases refer to Britain; Smith laments the "monopoly which our manufacturers have obtained against us" (WN 471).

⁶ In the previously discussed passage from WN IV.ii that anticipates the invisible hand, Smith himself exaggerates the public benefit that "every" investor brings. We read that every individual is continually striving to discern "the most advantageous" employment for his

words “everyone” and “everybody” never appear. The chapter ends, moreover, with Smith lamenting that the “private interests of many individuals”—along with “the prejudices of the publick”—constitute an insuperable obstacle to “the freedom of trade” being fully “restored” in Great Britain (*WN* 471).

Grampp returns a step toward reality later when he states that the merchant who keeps his capital at home promotes the “interest of everyone” because “domestic wealth is a resource on which the nation can draw to defend itself” (450). The exaggeration remains—aren’t there usually *some* inhabitants in a society whose “interests” are promoted when it is *less* able to defend itself?—but Grampp’s emphasis on national defense can remind us of a more important point. Nations often wield their military strength to devastate *foreigners*.

“THE ORDINARY REVOLUTIONS OF WAR AND GOVERNMENT”

Although Grampp ignores the destructiveness of war when he repeatedly invokes the benefit the invisible hand brings to “everyone,” military considerations are (as noted above) central to his argument. He elaborates that the individual profiled in the invisible-hand paragraph *would* understand how keeping his capital at home boosts domestic employment and output. The consequence the capital-owner would *not* fathom is the possible augmentation of his nation’s *power* (Grampp 454).⁷ How does

capital, and that his quest to promote his *own* advantage “necessarily” directs him to the employment that is “most advantageous” to the society (*WN* 454). Let me suggest a dramatic contemporary counterexample. If a methamphetamine dealer earns a windfall by hatching brilliant new techniques for production and distribution, does his contribution to the proliferation of “crank” addicts constitute a major contribution to American society? On *WN*’s tendency to deploy terms such as advantageous, proper, improved, interest, greatness, and justice in a materialistic or “economistic” fashion, see Minowitz 1993, 15-17, 34, 37-39, 46.

⁷ I feel compelled to point out that *WN*’s invisible-hand paragraph refers only once to what the agent *knows*, but four times to what he *intends*—and once to his intention. Grampp similarly stumbles later when he implies that the invisible hand has only one unintended consequence, “to contribute to the defense of the nation” (Grampp 446). Even if the benefit to domestic employment is easy to *know*, that benefit is also *unintended*, and Smith does not

Grampp make military power so important, given the absence of any reference to military affairs in the passages from *Wealth of Nations* we have examined?

One key premise is the claim, issued later in the invisible-hand chapter, that defense is “of much more importance than opulence” (*WN* 464-5), which Smith provides in defending trade-restrictions that promote an industry “necessary for the defence of the country.” Smith here defends the Navigation Act, which, although economically harmful, boosted the number of Britain’s sailors and ships (463); in his later chapter on government’s expenses/duties (V.i), Smith emphasizes that “the great expence of firearms gives an evident advantage to the nation which can best afford that expense” (708) and he laments the decay of “martial spirit” in commercial societies like Britain. To these passages (and others like them),⁸ Grampp adds considerations he admits are only inferences. Apparently drawing on the invisible-hand paragraph’s invocation of the capital owner’s “security,” Grampp infers that domestic capital is more “secure” than capital held abroad because it can more easily or reliably be marshaled to “support” defense (by funding military expenditures, I presume).⁹ But when Smith in his earlier chapter on “the natural progress of opulence” describes the differences in security among capital invested in land (highest), manufacturing (middle), and foreign commerce (lowest), his focus is on the situation of the *owner*, not the nation.¹⁰

Although Grampp may here go astray by confounding the nation’s security with the merchant’s, he is on much firmer ground when he invokes the grim conclusion of *Wealth of Nations*, Book III (Grampp 459). Smith here says that the capital “acquired to any country” via either manufacturing or foreign commerce is a “very precarious and uncertain possession” until

value it merely as a prop to national defense. Smith’s emphasis in IV.ii on unintended consequences figures prominently in many of the nine interpretations Grampp attacks.

⁸ Grampp usefully cites Smith’s claim “the great object of the political oeconomy of every country, is to encrease the riches and the power of that country” (*WN* 372). Also crucial are Smith’s statement that defense is “the first duty of the sovereign” (689) and his incorporation of societal “greatness” (along with wealth) within “the great purpose” that every political economy “system” intends to promote (687).

⁹ Grampp also hypothesizes that boosting domestic employment promotes national defense because workers abroad would be harder to summon for military service (Grampp 453).

¹⁰ *WN* 377-79. The capital of the landlord is “fixed in the improvement of his land” and “seems to be as well secured as the nature of human affairs can admit of” (*WN* 378); the “planter who cultivates his own land...is really a master, and *independent of all the world*” (emphasis added); the capital of the manufacturer, “being at all times within his view and command, is more secure than that of the foreign merchant” (379).

part of it has been “secured and realized in the cultivation and improvement of its lands” (WN 426). As Grampp highlights, Smith’s focus here is on *national* security; Smith proceeds to remind his readers that a merchant “is not necessarily the citizen of any particular country” and to assert that “a very trifling disgust” will cause a merchant to move his capital (and the industry it supports) from “one country to another.” The “ordinary revolutions of war and government easily dry up the sources of that wealth which arises from commerce only.” Yet even the “more solid improvements of agriculture” can be destroyed, as happened during the fall of the Roman Empire, by “a century or two” of barbarian depredations. The development of firearms ameliorates this danger, but leaves others in its wake.¹¹

I concede that it is easy to overlook some of the striking claims Smith makes on behalf of national-security issues, and that Grampp provides a major service by arguing for the connection between defense and the invisible hand. But if, as Grampp asserts, “the leading proposition of Smith’s economic policy” is that “defense is more important than wealth” (Grampp 442), why didn’t Smith title his book, *An Inquiry Into the Nature and Causes of the Defence of Nations*? If his main focus had been on military power, why would Smith offer his knowledge to all “nations” indiscriminately? It is possible, albeit unlikely, that most nations could be well defended, but military “power” also includes offensive capabilities; and millions of people have believed that economic liberty as touted by Smith serves to benefit some nations at the expense of other nations. Smith concedes that although “the wealth of a neighbouring nation” is “certainly advantageous in trade,” it is “dangerous in war and politicks” (WN 494). In Book IV, Smith persistently attacks what he alleges are the zero-sum aspects of mercantilism—its agendas for imperialism and colonization (588, 613, 626-7), its obsession with self-sufficiency (435, 456-7, 458, 493, 538-9) and “the balance of trade” (431-2, 450, 488-9, 642), its appeals to “national prejudice” and “national animosity” (474, 475, 494, 495, 496, 503), and its premise that trading nations advance their “interest” by “beggarizing all their neighbours” (493). His alternative is the “freedom of trade” (433, 464, 469, 580) that would allow many nations, if not all, collectively to advance “the accommodation and conveniency of the species” (30) and “the business of mankind” (592) via “the mutual communication of knowledge and of all sorts of improvements which an extensive commerce from all countries to all countries naturally, or rather necessarily, carries along with it” (627). He

¹¹ See Minowitz 1989.

asserts, perhaps implausibly, that foreign trade is continually occupied in performing “great and important services” and providing “great benefit” to *all* of the participating countries (447). He once even describes the typical smuggler as a man who “would have been in every respect, an excellent citizen, had not the laws of his country made that a crime which nature never meant to be so” (898).

Departing from Grampp, most scholars would locate “the leading proposition of Smith’s economic policy” at the conclusion of Book IV. In here providing his most complete overview of the “system of natural liberty,” Smith proclaims that “no human wisdom or knowledge could ever be sufficient” to provide the “sovereign” with the capability of “superintending the industry of private people” and “directing it towards the employments most suitable to the interest of the society” (*WN* 687-8). For Grampp, Smith uses the invisible hand to discourage governments from trying to prevent merchants from investing their *capital* abroad. But Smith’s reference here to “the industry of private people” should remind us that Smith also vigorously tried to discourage governments from “directing” the allocation of *labor*. The following passage is particularly vivid.

The patrimony of a poor man lies in the strength and dexterity of his hands; and to hinder him from employing this strength and dexterity in what manner he thinks proper without injury to his neighbour, is a plain violation of this most sacred property. It is a manifest encroachment upon the just liberty both of the workman, and of those who might be disposed to employ him.... The affected anxiety of the law-giver lest they should employ an improper person, is evidently as impertinent as it is oppressive.¹² (*WN* 138)

According to Smith, the system of natural liberty would have a dramatic impact in harnessing “[t]he natural effort of *every* individual to better his

¹² Between two passages that tout the liberty of colonists “to manage their own affairs their own way” (*WN* 572, 584), Smith invokes “the most sacred rights of mankind” to condemn policies that prohibit “a great *people*” from “employing their stock and *industry* in the way that they judge most advantageous to themselves” (582; emphasis added). Also relevant are his enthusiasm for “the free circulation of labour” (135) and his criticism of institutions or policies that obstruct it: “exclusive corporations” (146), apprenticeships (151) and the Poor Laws (152).

own condition” (540), indeed, “the natural effort which *every* man is *continually* making to better his own condition” (674) [emphasis added]—not just the natural effort of merchants involved in foreign trade.

“THE ECONOMY OF GREATNESS”

Although the invisible hand surfaces only once in *Wealth of Nations*, the book is pervaded by the prospect of an unseen agency—perhaps an unseen intelligence—that constructively channels the behavior of self-interested individuals and should deter political elites from being overly intrusive. In passages I discuss above—and in countless others—Smith invokes *nature* as the principle or authority to which such leaders should defer.¹³ To the hordes who condemn Smith for speaking of “natural” liberty—and especially for painting it in such an optimistic light—*The Theory of Moral Sentiments* might be even more objectionable because it portrays nature as exuding both power and benevolent purpose. *Moral Sentiments*, like *Wealth of Nations*, includes one reference to an invisible hand. Only in *Moral Sentiments*, however, does Smith attribute the invisible hand to Providence and speak frequently of nature’s “wisdom,” which he links with God. Only

¹³ Friedrich Hayek and libertarians who highlight “spontaneous order” typically refrain from invoking any sort of non-human authority or intelligence. Hayek credits Smith (and other 18th-century Scots) for showing that “an evident order which was not the product of a designing human intelligence need not therefore be ascribed to the design of a higher supernatural intelligence.” Because “no human mind can comprehend all the knowledge which guides the actions of society,” Hayek exhorts us to conceive of “an effective coordination of human activities without deliberate organization by a commanding intelligence”; such coordination often occurs via an “impersonal mechanism” such as a market (Hayek 1959, 4, 59, 159). Emphasizing the limits on the knowledge a human individual can attain, Hayek (like Smith) encourages his readers to assume “an attitude of humility towards the impersonal and anonymous social processes by which individuals help to create things greater than they know.” These “impersonal and anonymous” processes would include languages, markets, and a variety of laws and customs (Hayek 1948, 7, 8, 11, 15, 22, 32, 86-88). Hayek could complain that Smith’s appeals to an invisible hand, “the wisdom of nature” (*WN* 674), and so on, may encourage readers to mis-identify impersonal social processes as a superhuman intelligence that leads or directs us. Although *TMS* goes further with its frequent appeals to a superhuman *designer*, it anticipates Hayek by explaining how moral consciousness and conduct can emerge via the purely human interactions that create “the impartial spectator.” For a penetrating discussion of Hayek in connection with Smith’s invisible hand, see Rothschild 2001, 140-2, 145-53, 155.

in this book does Smith invite the reader to imagine an invisible *hand* that fulfills the intentions of a superhuman being—and that shows particular care for the poor. Only in this book does Smith hint that people will be neither happy nor moral unless they believe in an afterlife (*TMS* 120-1, 131-2, 164), and only here does Smith ridicule “power and riches” as “trinkets of frivolous utility” (181-2). Grampp, alas, fails to convey these momentous contrasts between the two books—and he misreads the paragraph that presents the invisible hand.

In treating the *Moral Sentiments* invisible hand, Grampp does accurately recount the starting point. Smith is arguing that mankind has consistently survived and progressed despite pronounced inequality. A “proud and unfeeling landlord” may exult in his ownership of “extensive fields,” but he cannot *eat* any more of the produce than can “the meanest peasant.” Smith proceeds to argue that the soil “maintains at all times nearly that number of inhabitants which it is capable of maintaining.” Shifting his attention from the landlord, Smith claims that “the rich” get to eat *better*, but not much *more*, than the poor eat;¹⁴ despite their “natural selfishness and rapacity” and their “vain and insatiable desires,” the rich end up sharing.

They are led by an invisible hand to make nearly the same distribution of the necessities of life which would have been made had the earth been divided into equal portions among all its inhabitants; and thus, without intending it, without knowing it, advance the interest of the society and afford means to the multiplication of the species. (*TMS* 184-85)

Grampp acknowledges key similarities between this invisible hand and the one in *Wealth of Nations*—each has a “favorable connotation,” presumably because each “leads the selfish to help others and to help them without a cost to themselves” (Grampp 463). He is right to challenge the plausibility of the *Moral Sentiments* version, but he ignores the disturbing lessons suggested by the surrounding material, and he goes embarrassingly astray in laying out the particulars.

¹⁴ However difficult it would have been for Smith to prove this thesis when he wrote, it would be harder for someone today to argue that the soil “maintains at all times nearly that number of inhabitants which it is capable of maintaining.” Millions are obese, while millions are starving. In any case, Smith thrice in the invisible-hand paragraph places great weight upon the adverb “nearly.”

When he attempts to specify the effects the invisible hand has on the rich, Grampp offers a fantasy.

They imagine there is no limit to what they can enjoy and so order whole harvests to be brought to them. They then discover “the eye is larger than the belly” and must find something to do with what they cannot use themselves. And what is it? They give it to the poor. . . . (Grampp 463)

For Grampp, this invisible hand thus differs from the *Wealth of Nations* hand because the relevant self-interest calls to mind “dumbbells who buy more than they can use and find themselves giving away much of it.” The stupid landlords, furthermore, “never learn”—otherwise “there would be only one redistribution,” after which “there would be no leftovers for the poor” and the invisible hand’s work would be done (Grampp 463).

If Grampp had scrutinized merely the paragraph in which the invisible hand appears, he could have provided a far superior elaboration. Smith states clearly that the landlord distributes the surplus food to the people who *prepare* the food “he himself makes use of,” to those who “fit up the palace” in which he dines, and to anyone else who provides or maintains “all the different baubles and trinkets, which are employed in the oeconomy of greatness” (*TMS* 184). Whether the relevant non-landlords are workers, servants, serfs, slaves, offspring, or wives, the reader confronts an ongoing “oeconomy”—a word that Smith rarely uses in *Moral Sentiments*—of *exchange*, not a one-time *gift* from a dim-witted landlord who initially thought he could consume the entire produce of his land.¹⁵ The reader also encounters an invisible hand that advances, via the “natural selfishness” of various individuals, “the interest of the society” and the propagation of the species—an invisible hand that harmonizes with most of the broad interpretations of *Wealth of Nations* that Grampp is criticizing.

As we have seen, Grampp lambastes *Moral Sentiments* partly because of his inference that the landlords are idiots who keep biting off more than they can chew and then disgorging the residue. Grampp and countless

¹⁵ In his final paragraph on *TMS*, Grampp admits the implausibility of thinking that the rich, in Smith’s account, simply “gave away much of their income.” Thus, Smith is “said to have meant that they [the rich] help the poor by giving them employment.” If this were true, Grampp adds, the poor would “get their income from working, not from leftovers, and an invisible hand is not needed to explain that” (Grampp 463). But *TMS* does bring employment clearly into the picture, and Grampp himself emphasizes the effects the invisible hand of *WN* has on employment.

other readers, furthermore, are skeptical about Smith's claims that the distribution of "the necessities of life" is "nearly" the same as it would have been if the earth were "divided into equal portions among all its inhabitants"—and that the soil *at all times* maintains "nearly that number of inhabitants which it is capable of maintaining." So let us dig deeper into the chronological foci of Smith's account.

The remarks quoted and paraphrased above are all in the present tense: the landlord "views" his large fields and the rich "select" the choicest produce, while the poor "derive" all that they need to subsist. The paragraph also begins in the present tense: "And it is well that nature imposes upon us in this manner" (*TMS* 183). To fathom *this* claim, however, we must address profound issues that Grampp's article ignores—and that Smith scholarship often depreciates.

Two paragraphs earlier in this short chapter (Part IV, chapter 1), Smith sketches the tragic fate of "[t]he poor man's son, whom heaven in its anger has visited with ambition." Abandoning the "real tranquility" that was "at all times in his power," the son endures a lifetime of study, toil, fatigue, worry, obsequiousness, and betrayal. As death approaches, he finally learns that wealth and greatness are "mere trinkets of frivolous utility, no more adapted for procuring ease of body or tranquility of mind" than the tweezers-cases lugged around by "the lover of toys" (*TMS* 181). Smith now broadens his focus to explain why the palaces, gardens, equipage, and retinue of "the great" stir up universal longing. Despite their frivolity, such trinkets captivate us because "that love of distinction so natural to man" is readily augmented by our tendency to become infatuated by the potency of the *things* (tools, machines, and "systems") that help us gratify our wishes (182). Smith then expands the lesson he drew from the parable of the poor man's son. When a person's vanity is eclipsed by "the languor of disease and the weariness of old age," or when he is compelled by "either spleen or disease to observe with attention his own situation," power and riches will finally appear to be "what they are," namely:

Enormous and operose machines contrived to produce a few trifling conveniencies to the body, consisting of springs the most nice and delicate, which must be kept in order with the most anxious attention, and which in spite of all our care are ready every moment to burst into pieces, and to crush in their ruins their unfortunate possessor.... They keep off the summer shower, not the winter storm, but leave him always as much, and sometimes more,

exposed than before, to anxiety, to fear, and to sorrow; to diseases, to danger, and to death.¹⁶ (182-3)

It must be emphasized that the two quasi-synonymous pairs of general terms that Smith here impugns—*wealth and greatness* as “trinkets of frivolous utility,” and *riches and power* as “operose machines” that perpetually threaten to destroy their “unfortunate possessor”—are precisely the pairs that *Wealth of Nations* deploys to identify the “object” or “purpose” of political economy (WN 372, 687).¹⁷ And political economy is the scientific genre into which Smith places *Wealth of Nations*.¹⁸ In light of these and other complexities, Grampp deserves praise for accentuating the evasions and enigmas that help define Smith’s legacy (Grampp 442, 455, 462-4).¹⁹

¹⁶ Whereas Grampp imagines moronic landlords who never learn that the eye is larger than the belly (463), Smith chides the “poor man’s son”—and “our conduct” generally (TMS 181)—for repeatedly forgetting that the “machines” that protect us from the summer shower are helpless against “the winter storm” (181-83). Smith also laments the loss in leisure, ease, and “careless security” caused by our vanity-inspired quest for wealth and power at TMS 50-51. On vanity’s contribution to the ubiquitous drive for “bettering our condition,” compare TMS 50-51 with WN 190, 341-42, and 869-70.

¹⁷ The later passage (WN 687)—which asserts that every “system” of preference or restraint ends up subverting “the great purpose which it means to promote.... the progress of the society toward real wealth and greatness”—does not mention political economy, but the term is strongly implied. The title of the relevant Book (IV) is “Of Systems of political Oeconomy,” which highlights the mercantilist and agriculturalist (e.g., Physiocratic) approaches as the political economy “systems” marred by preferences and/or restraints (Smith introduces the “system of natural liberty” at the end of Book IV). On p. 372, in any case, Smith proclaims that “the great object of the political oeconomy of every country, is to encrease the riches and power of that country.”

¹⁸ When Smith speaks of “what is properly called” political economy, he uses language that specifies the subject matter of his world-renowned book: “the nature and causes of the wealth of nations” (WN 678-79). WN’s title does *not* mention greatness or power, and its text spends relatively little time defining or discussing them. Another prominent definition likewise elevates wealth/riches above greatness/power: in the brief introduction to Book IV, Smith explains that political economy, “considered as a branch of the science of a statesman or legislator, proposes....to enrich both the people and the sovereign” (428). For a sketch of how WN addresses the relationship between wealth/riches and greatness/power, see the “Ordinary Revolutions” section above.

¹⁹ As Grampp puts it, “[t]he effort to reconcile the diverse ideas is the greatest of the efforts a reader must make in order to understand the *Wealth of Nations*, greater certainly than the effort needed to understand a particular idea when it is taken by itself” (Grampp 460-1). In Minowitz 1993, I challenge the dominant trends in contemporary scholarship on Smith and strive to reopen the “Adam Smith Problem” posed by the contrasts between his two books. Individuals interested in the formidable complexity of Smith’s writing and thinking should, at a minimum, consult the recent books by historian Jerry Muller (1993), economist Vivienne Brown (1994), and philosophy professor Charles Griswold (1999).

One cannot resolve the trinkets conundrum by assuming that Smith underwent an epiphany after 1759, when the first edition of *Moral Sentiments* appeared. At the start of a chapter (I.iii.3) added for this work's final edition in 1790, fourteen years after the publication of *Wealth of Nations*, Smith wrote that the disposition to admire wealth and greatness is "the great and most universal cause of the corruption of our moral sentiments" (TMS 61).

Smith's depiction of wealth and greatness as trinkets becomes even more complex in the paragraph that follows the one that ridicules the "[e]normous and operose machines" and that immediately precedes the paragraph on the invisible hand. Smith associates his denunciation of wealth and greatness with a "splenetic philosophy," familiar to everyone in times of "sickness or low spirits," that views things in an "abstract and philosophical light." But he proceeds to say that the same objects—when we view them from the more "complex" perspective that emerges in times of ease and prosperity—will appear "grand," "beautiful," and "noble," and hence as worthy of "all the toil and anxiety" we typically bestow upon them (TMS 183). Smith has provided clues, but he never directly mediates between the two competing perspectives: sick/old/philosophical versus healthy/young/prosperous.

The invisible-hand paragraph opens in the present tense: "it is well that nature imposes upon us in this manner." Smith labels the above-described infatuation with systems and machines a "deception," but lauds it because it "rouses and keeps in continual motion the industry of mankind" (TMS 183); Smith here speaks about phenomena that are contemporaneous to him (as he does a few sentences later when he discusses the landlord's fields and the invisible hand that assists the poor). However, he immediately shifts to a retrospective view as he celebrates the deception as the spring of human progress.

It is this which first prompted them to cultivate the ground, to build houses, to found cities and commonwealths, and to invent and improve all the sciences and arts, which ennoble and embellish human life; which have entirely changed the whole face of the globe, have turned the rude forests of nature into agreeable and fertile plains, and made the trackless and barren ocean a new fund of subsistence, and the great high road of communication to the different nations of the earth. (TMS 183-84)

After adding the claim that the earth “by these labours of mankind has been obliged to redouble her natural fertility, and to maintain a greater multitude of inhabitants” (184), Smith presents the invisible-hand scenario about the “proud and unfeeling landlord.”

Let me summarize. Our population has grown because “nature” tricked us into laboring that transforms the earth, partly by multiplying the earth’s “natural” fertility. The invisible hand serves to maintain “the multiplication of the species” in the face of widespread landlessness. Under both scenarios, we advance collectively despite two types of moral shortcomings: the selfishness, rapacity, callousness, vanity, and pride that tarnish the economic elite (landlords and “the rich”); and the “natural” and widespread “love of distinction” that can prompt even “the poor man’s son” to sacrifice tranquility and happiness in the frivolous pursuit of “trinkets” (181-82). Nature wields its power and achieves its ends in complex if not paradoxical ways. Adam Smith grasps the two disparate perspectives on wealth and greatness: the “splenetic” negative perspective and the “complex” positive one. Unlike the rich, he cares for the poor; unlike most of us (including the poor man’s son afflicted by ambition), he is never intoxicated by the “trinkets of frivolous utility.”²⁰

Smith’s contribution is philosophical, one may infer, since he fathoms the paradoxical truths about how everything fits together. His

²⁰ Is this *TMS* chapter the work of a “tyro,” as Grampp suggests (463), or the work of a sage? Seventy-three years before the *Journal of Political Economy* published Grampp’s article, it published “Adam Smith and Laissez Faire,” a pioneering article by Jacob Viner. Viner skillfully displays the theological clash (and some related differences) between Smith’s two books, but exaggerates both the optimism and the dogmatism in *TMS*. In faulting *TMS* for “absolutism” and “rigidity” (Viner 1958, 216), Viner ignores the complex dialectics of the trinkets puzzle. He likewise overstates the extent to which *TMS* posits “universal and perfect harmony” and presents an “unqualified doctrine of a harmonious order of nature, under divine guidance” (217, 220, 222-23). Viner overlooks mankind’s continual vulnerability to “the winter storm,” anxiety, fear, sorrow, disease, danger, and death (*TMS* 183). In addition, he overemphasizes the passages (*TMS* 105, 166, 168) that identify the *happiness* and *perfection* of “the world” and its “species” as the purposes of Nature/God (Viner 1958, 217, 220, 229-30); and he ignores the passages that highlight individual preservation and species propagation (*TMS* 77, 87), humbler goods that resemble *WN*’s humbler articulation of “the wisdom of nature” (*WN* 673-74; see pages 408-409 below). Like Grampp, finally, Viner is too quick to invoke Smith’s alleged “absentmindedness” to explain inconsistencies (Viner 1958, 241). Anyone who savors the delicacy of Smith’s prose, and tracks the multitude of minute changes Smith made in revising his two books, has no reason to doubt his 1788 description of his approach as an author: “I am a slow a very slow workman, who do and undo everything I write at least a half a dozen of times” (Letter to Thomas Cadell, 15 March 1788). The above Viner citations are from the reprinting in a book by Viner (1958).

contribution is also rhetorical. By arguing that we are “led”—certainly some of the time, perhaps most of the time—by an invisible hand to ends we did not intend to promote, Smith reminds us that we are supreme in neither comprehension nor power. At the conclusion of IV.1, however, Smith does smile on certain efforts to promote broad public benefits. It turns out that our “love of system”—our attraction to the “beauty of order, art and contrivance,” the attraction that helps wealth and greatness seduce us—can fruitfully be manipulated to “implant public virtue in the breast of him who seems heedless of the interest of his country.” To do this, you could proceed by describing “the great system” of public policy that helps feed, clothe, and house “the subjects of a well-governed state.” After explaining “the connections and dependencies of its several parts...and their general subservience to the happiness of the society,” you could “show how this system might be introduced into his own country,” describing the current “obstructions” and how they might be removed so that “the wheels of the machine of government” would “move with more harmony and smoothness” (*TMS* 185-86). From Smith’s point of view, obviously, *Wealth of Nations* is well suited to “implant public virtue” along these lines. But this book also calls upon the invisible hand, and many powerful arguments, to inoculate kings, princes, legislators, and statesmen from the “innumerable delusions” that would afflict anyone who sought to superintend the “industry of private people” (*WN* 687).

A similar warning, which particularly seems to challenge Part IV’s suggestions about using the “love of system” to bolster civic virtue, suffuses some passages in Part VI of *Moral Sentiments*, which Smith added for the 1790 edition. People “intoxicated by the imaginary beauty” of an “ideal system,” Smith now warns, often succumb to “the madness of fanaticism” (*TMS* 232). The “man of system” who ignores “the great interests” or “strong prejudices” that may oppose his “ideal plan of government,” furthermore, treats people as “the hand arranges the different pieces upon a chess-board” (note the impact of a *visible* hand). Such a man fails to recognize that “in the great chess-board of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might chuse [*sic*] to impress upon it” (234).

“A FEW LORDLY MASTERS”

By invoking an invisible hand to drive home human shortcomings in power, wisdom, and virtue, Part IV of *Moral Sentiments* communicates a lesson that most religions emphasize. And in the sentence after the one that describes the invisible hand, Smith incorporates a divine presence missing from *Wealth of Nations*. As he did in the preceding sentences, Smith reassures his readers about the fate of the masses deprived of land (and power).

When Providence divided the earth among a few lordly masters, it neither forgot nor abandoned those who seemed to have been left out in the partition.... In what constitutes the real happiness of human life, they are in no respect inferior to those who would seem so much above them. In ease of body and peace of mind, all the different ranks of life are nearly upon a level, and the beggar, who suns himself by the side of the highway, possesses that security which kings are fighting for. (*TMS* 185)

Although Smith in *Wealth of Nations* does offer a friendly comment on “the Deity” that ancient Greek physicists investigated as a “part” of “the great system of the universe” (*WN* 770)—and a disparaging comment on the superstitious recourse to “gods” (767)—he never mentions God or Providence, and he portrays nature in a less exalted light. His grimmer posture toward the cosmos corresponds to his harsher accounts of starvation and land ownership. Regarding starvation, the Introduction laments the plight of primitive “nations” that subsist via hunting and fishing. Even though almost every able-bodied person works, these societies are so poor that they sometimes are forced to kill infants, old folks, and people “afflicted with lingering diseases”—or to abandon such individuals “to perish with hunger, or to be devoured by wild beasts.” In “civilized and thriving nations,” by contrast, “all are often abundantly supplied” despite the “great number” of persons who consume lavishly even though they do not work (10).

In *Moral Sentiments*, Smith lauds the invisible hand of Providence for ensuring, in all times and places, that the human “species” survives and multiplies. *Wealth of Nations* proceeds in a far more empirical fashion. Smith depicts both starvation and famine. As in *Moral Sentiments*, however, Smith

does *not* place the blame on the monopolization of land ownership by “a few lordly masters” (*TMS* 185). Hunger and mortality plague hunting/fishing societies, despite their egalitarian economic arrangements—there simply is no property that “exceeds the value of two or three days labour” (*WN* 709) and the “[u]niversal poverty establishes...universal equality” (712). Circumstances improve as society advances “naturally” into the three subsequent “periods” or “states”: herding/pasturage, agriculture, and commerce (trade and manufacturing). But the torments of our origins recur even in the last two stages.²¹

In his most detailed discussion of food shortages, Smith focuses on the experience of Europe during recent centuries. He concedes that “dearths” have arisen from “real scarcity” caused sometimes by “the waste of war” but more often by “the fault of the seasons”; such scarcity can be ameliorated but not eliminated (*WN* 526-7). By blaming the seasons for dearths, Smith is blaming nature. Famine, on the other hand, “has never arisen from any other cause but the violence of government attempting, by improper means, to remedy the inconveniences of a dearth” (526). By tracing famines to abusive governments, Smith paves the way for nature’s remedy—the “unlimited, unrestrained freedom of the corn trade”—which is also the “best palliative” of dearths (527; cf. 538).

When he discusses subsistence and propagation in general terms, beyond the current situation in Europe, Smith likewise leaves us with questions about how nature and human institutions interact. One dilemma society confronts is that, as “[e]very species of animals naturally multiplies in proportion to the means of their subsistence” (*WN* 97), prosperity causes childhood mortality to decrease, which eventually causes wages to decrease. In a stationary economy, the “great body of the people” merely subsist; in a decaying economy they die off (86-8, 90-1, 97-9). Smith suggests China as an example of the stationary state. It “has long been one of the richest, that is, one of the most fertile, best cultivated, most industrious, and most populous countries in the world”; yet centuries before Smith’s time, it had “perhaps...acquired that full complement of riches which the nature of its

²¹ The four-stages theory is infused by something like an invisible hand insofar as Smith says nothing to suggest that human leaders or visionaries have played, or are needed to play, a role in propelling society from one stage to the next (cf. *WN* 422 on the “great revolution” that brought down feudalism). Needless to say, none of the four stages involves human fulfillment of a *divine* plan.

laws and institutions permits it to acquire.”²² In *all* of its “great towns,” tragically, children are “every night exposed in the street, or drowned like puppies in the water.” Furthermore, for the hundreds (perhaps thousands) of underfed people in Canton who live on rivers and canals in fishing boats—and are “eager to fish up the nastiest garbage” thrown overboard from a European ship—a putrid cat carcass is “as welcome...as the most wholesome food to the people of other countries” (89-90).²³ Do *these* landless beggars sun themselves on the banks of the river and enjoy “that security which Kings are fighting for”? Does the invisible hand of Providence bring them “ease of body and peace of mind”? Did Smith ever really believe that “all” of “the works of nature” were intended to promote “[t]he happiness of mankind” and to “guard against misery” (*TMS* 166)?²⁴

The evolution of society beyond the hunting stage also introduces threats to the economically *advantaged*. Smith describes, in stark terms, the plight of the owner of valuable property acquired by the “labour of many years”: he is “at all times surrounded by unknown enemies...from whose injustice he can be protected only by the powerful arm of the civil magistrate” (*WN* 710).

Obviously, reading aloud the *Moral Sentiments* passage extolling the “real happiness” enjoyed by the beggar cannot typically neutralize the dangers economic inequality poses. Only in *Wealth of Nations* does Smith

²² In the next chapter, Smith speaks more confidently: it is “probably” (not *perhaps*) the case that China had acquired all the riches it could, given “the nature of its laws and institutions” (*WN* 111). Smith proceeds to elaborate the toll exacted by those laws and institutions, particularly the obstacles to foreign commerce and the vulnerability of the poor and “owners of small capitals” to being “pillaged and plundered” by public officials (111-12, 680-81).

²³ We can only imagine the “[w]ant, famine, and mortality” that would afflict the beggars in a *shrinking* economy, “where the funds destined for the maintenance of labour were sensibly decaying.” Smith suggests that this condition may obtain in some of Britain’s colonies in India (*WN* 90-1), and later elaborates the pernicious policies of the East India Company (635-41, 751-53).

²⁴ Unlike Viner (and others), I am not prepared to belittle *TMS* as a juvenile work. Viner asserts that Smith, when he wrote this book, was a “purely speculative philosopher, reasoning from notions masquerading as self-evident verities” (Viner 1958, 230). Viner here overlooks the empirical components of *TMS*—e.g., the way Smith uses “sympathy” and “the impartial spectator” to explain how moral standards and behavior emerge from widespread patterns of human interaction—many of which remain plausible. Regarding *WN*, however, Viner is wise to suggest that statements about natural harmony may be “obiter dicta, thrown in as supernumerary reinforcements of an argument already sufficiently fortified by more specific and immediate data” (Viner 1958, 224). For a *Journal of Political Economy* article that does justice to *TMS* (and to Smith’s philosophical essays), see Bitterman 1940. Particularly valuable are Bitterman’s elaboration of the Newtonian aspects of Smith’s approach (497-504, 511-16, 717).

provide detailed explanations of how sustenance can trickle down from wealthy owners of land and capital. Consider first the herding stage: a “Tartar chief, the increase of whose herds and flocks is sufficient to maintain a thousand men,” cannot exchange his surplus “rude produce” for “any manufactured produce, any trinkets and baubles.” He therefore employs the surplus by “maintaining a thousand men,” who in exchange can provide only obedience; the chief’s authority becomes “altogether despotical” (*WN* 712-13).²⁵

In its early moments, the agricultural stage features shepherd-like political arrangements: the “sovereign or chief” is simply “the greatest landlord of the country.” One example is “our German and Scythian ancestors when they first settled upon the ruins of the western empire” (*WN* 717). Smith elaborates this earlier, in Book III, where he provides his most detailed discussion of the relationship between lords and their subordinates. A “great proprietor” in feudal Europe, lacking access to foreign commerce and “the finer manufactures,” consumed his entire surplus in “rustick hospitality” that in effect purchased the allegiance of servants along with a “multitude of retainers and dependents” (413-14). The feudal proprietor thus resembles the shepherd chief.

According to *Moral Sentiments*, Providence “divided the earth among a few lordly masters.” This description could not apply to the hunting stage, as presented by *Wealth of Nations*, for two reasons: there are no lordly masters who own the land, and widespread poverty inhibits “the multiplication of the species” (*TMS* 184-85). As we have seen, however, the description does apply to the herding stage—except that the “masters” here monopolize herds rather than fixed tracts of land.²⁶ And the description

²⁵ Such a “little sovereign” ends up being supported by “a sort of little nobility”: “Men of inferior wealth combine to defend those of superior wealth in the possession of their property, in order that men of superior wealth may combine to defend them in the possession of theirs” (*WN* 715).

²⁶ *WN* also refers to “the original state of things, which precedes both the appropriation of land and the accumulation of stock”; here a laborer is not required to share his produce with either “landlord or master” (*WN* 82). The hunting/fishing stage seems to fit these criteria; in the second stage, the “chief” controls the herds and their produce; in the hunting stage, there is “little or no authority or subordination” (712-13). When Smith states that the tiller of the soil “generally” has his maintenance “advanced to him from the stock of a master, the farmer who employs him,” Smith seems to be describing the final two stages. Workers in “all arts and manufactures,” similarly, usually need a “master” to advance them “the materials of their work, and their wages and maintenance till it be completed” (83). The majority of human beings, except among hunting/fishing societies in which harsh poverty is universal, are thus subject to economic “masters,” and Providence is not responsible.

applies to feudal arrangements in Europe that more or less represent the agricultural stage. But when Smith describes the origins of feudalism, he offers a cynical explanation that invokes neither nature nor Providence: “the chiefs and principal leaders” of the conquering Germans and Scythians simply “acquired or usurped to themselves the greater part of the lands of those countries” (*WN* 381-82). For many years thereafter, “the open country” in Europe was a “scene of violence, rapine, and disorder” (418).

Nature and convention also interact complexly in Smith’s account of primogeniture and entails, institutions that in effect helped certain lordly masters to maintain monopolistic patterns of land ownership in Europe. Under feudal conditions, primogeniture and entails “might not be unreasonable,” since large estates supported political authority in “those disorderly times.” Sustained by family pride even in Smith’s day, however, primogeniture and entails remained major obstacles to the subdivision and commercialization necessary for full agricultural development (*WN* 382-86). Primogeniture and entails surely belong among the “human institutions” that the preceding chapter blamed for having “disturbed the natural course of things” in Europe (377-78) and having “inverted” what the chapter title (III.i) identifies as “the natural progress of opulence.”²⁷

Smith’s account in Book III of the demise of feudalism and the emergence of commercial society draws on elements of both invisible hands (*WN* IV.ii and *TMS* IV.1). The power of the lords was “gradually” ended by “the silent and insensible operation of foreign commerce and manufactures.” Whether or not a hand can be invisible, it can certainly be silent. There are much stronger echoes, in any case. Once the lords had the chance to purchase “frivolous and useless” items (e.g., diamond buckles) that could be “all their own,” they lost their disposition to “share” their surplus, and thus “gradually bartered their whole power and authority” for the sake of “the most childish, the meanest and the most sordid of all vanities” (*WN* 418-19).²⁸ After completing the story by explaining how the

²⁷ On the prominence of family pride and inherited wealth in sustaining shepherd-stage authority generally, see *WN* 714 and 421-22. Yet another important perspective on land-ownership patterns emerges in Smith’s discussion of colonies, where he states that “[p]lenty of good land, and liberty to manage their own affairs their own way, seem to be the two great causes of the prosperity of all new colonies” (*WN* 572; cf. 566-67, 570, 572, 584).

²⁸ In both books, Smith sometimes directs vicious criticism at the economically privileged. According to *TMS*, as we have seen, the landlord is “proud and unfeeling,” while “the rich” are characterized by “natural selfishness and rapacity” and “vain and insatiable desires” (*TMS* 184). Smith speaks similarly during *WN*’s discussion of the feudal lords who traded their authority for trinkets: “All for ourselves, and nothing for other people, seems, in every age of the world, to have been the vile maxim of the masters of mankind” (*WN* 418).

lords similarly allowed their tenant *farmers* to become independent, Smith observes that the “great proprietors” thus “sold their birth-right, not like Esau for a mess of pottage in time of hunger and necessity, but in the wantonness of plenty, for trinkets and baubles.”²⁹ Smith speaks even more generally in the following, widely cited passage.

A revolution of the greatest importance to the publick happiness, was in this manner brought about by two different orders of people, who had not the least intention to serve the publick. To gratify the most childish vanity was the sole motive of the great proprietors. The merchants and artificers, much less ridiculous, acted merely from a view to their own interest, and in pursuit of their own pedlar principle of turning a penny wherever a penny was to be got. Neither of them had either knowledge or foresight of that great revolution which the folly of the one, and the industry of the other, was gradually bringing about. (*WN* 422)

Grampp chides the scholars who see the invisible hand at work here, and is skeptical about whether we can specify the “relation” between the hand’s two versions (Grampp 464). To me, there are obvious connections involving globalization, the monopolization of land, the contribution “trinkets and baubles” make in promoting public benefit via private vice, and the complex dialectics that infuse Smith’s accounts of how nature shapes human morality, psychology, and institutions.³⁰

Smith seems to define commercial society in the following terms: “[e]very man lives by exchanging, or becomes in some measure a merchant” (*WN* 37). The fall of feudalism thus transformed rather than eliminated dependence. In modern Europe, each “tradesman or artificer derives his subsistence” from the employment of “a hundred or hundred thousand

²⁹ *WN* 421. Recall how Smith simply *identified* “manufactured produce” with “trinkets and baubles” when explaining that a shepherd chieftain could only employ his surplus by “maintaining” a multitude of subordinates (*WN* 712); and recall the prominence of “baubles and trinkets” in the invisible-hand paragraph of *TMS* (184).

³⁰ The two discussions, furthermore, are similarly located in their respective works: Book IV of *WN* and Part IV of *TMS* (*TMS* is divided into parts rather than books). The account of feudalism occupies the central book of *WN*, and is followed quickly by the invisible hand, which lies roughly in the middle of *WN*, page-wise. In *TMS*, similarly, the invisible hand appears in the central part.

different customers”; though he is partly “obliged to them all,” he is not “absolutely dependent” on any one (420); in a “civilized” society, the division of labor renders *everyone* dependent on “the assistance and cooperation of many thousands” (22-3, 26). Without intervention by government, furthermore, the division of labor threatens to annihilate the “intellectual, social, and martial virtues” among “the great body of the people” (781-2). Again, what became of the Providence that provides for the “real happiness” of the lowly?

Moral Sentiments does not employ the four-stages theory, although on several occasions it contrasts the harsh conditions of “savage” life with the ease of “civilized” circumstances.³¹ In a section Smith added for the 1790 edition, he does offer a remarkable generalization that calls to mind passages from *Wealth of Nations* about modern Europe: in “commercial countries,” the “authority of law is always perfectly sufficient to protect the meanest man in the state” (TMS 223).

Another dramatic echo of the invisible hand resonates in the sub-chapter of *Wealth of Nations* whose theme is religion. The medieval Church, Smith boldly suggests, was “the most formidable combination that ever was formed...against the liberty, reason, and happiness of mankind” (WN 802-3). It controlled large tracts of land; like the lords, it gained political authority by distributing agricultural surpluses (“profuse hospitality” and “extensive charity”). Its power surpassed that of the lords for two reasons: its temporal force was accentuated by “spiritual weapons” and “the grossest delusions of superstition;”³² and it could act as “a sort of spiritual army, dispersed in different quarters,” whose “movements and operations” were “directed by one head, and conducted upon one uniform plan” (800-3). This unprecedented “combination” eventually collapsed in the same way that the pernicious power of the barons did. Even though “all the wisdom and virtue of man” could never even have “shaken” it, nature—here, “the natural course of things”—again came to the rescue via the “gradual improvements of arts, manufactures, and commerce” (803). Contra Grampp,

³¹ In TMS’s most sustained discussion of the differences between primitive and civilized societies (Part V, Chapter 2), Smith condemns the infanticide practiced by “the polite and civilized Athenians” (TMS 210). Although WN’s Introduction eschews condemnation and portrays infanticide among hunting/fishing nations as a regrettable necessity (WN 10), TMS here—as elsewhere—conveys higher standards, saying only that infanticide is “undoubtedly more pardonable” in the “rudest and lowest” stage (TMS 210).

³² Recall how mercantilism drew on both public “prejudices” and private “interests” in sustaining itself (WN 471). Smith likewise links prejudices and interest in explaining his famous assertion equating the “laws concerning religion” with the “laws concerning corn” (539).

there are many reasons to think that Smith, in sketching the roots of modernity, incorporated some of the “many other cases” in which an invisible hand linked with commerce led a person to “promote” a beneficial “end which was no part of his intention” (*WN* 456).

“DESIGNING POWER”

Readers of Smith encounter a third invisible hand—“the invisible hand of Jupiter”—in a posthumously published essay that Grampp expounds insightfully but briefly (Grampp 461-2). As we have seen, Smith sometimes presents sweeping claims that he himself may have regarded as exaggerations; in comparing *Wealth of Nations* and *Moral Sentiments*, I have suggested that Smith resorts to other types of rhetorical maneuvering (especially regarding the character of the agency or intelligence that the invisible hand embodies). The third manifestation of the invisible hand raises another set of questions about the relationship between Smith’s two books, the way each of them blends science with rhetoric, and his posture toward religion. Smith bequeathed to the world a unique combination of lucid sentences and enigmatic books.³³

Prior to 1759, when *Moral Sentiments* was published, Smith drafted three essays about “the principles which lead and direct philosophical enquiries.”³⁴ The essay that is by far the longest addresses the history of astronomy. While discussing “the first ages of society,” Smith contemptuously invokes the “the invisible hand of Jupiter” to illustrate the “pusillanimous superstition which ascribes almost every unexpected event, to the arbitrary will of some designing though invisible beings.” Smith lists eclipses, thunder, lightning, comets, and meteors among the dazzling natural phenomena that people attributed to “intelligent, though invisible causes.” People experienced *themselves* taking actions that altered the external world, and therefore imagined that a *divine* agency or “designing power” was responsible for the “irregular events” that surprised them. But even the

³³ Smith freely deploys understatement as well as overstatement. I’ve been emphasizing his exaggerations, but the equivocations, insinuations, and qualifications (e.g., the ubiquitous “perhaps”) may be more prevalent. Cf. Viner 1958, 222-23 on *WN*’s recourse to phrases such as majority, frequently, “in most cases,” and “in general.”

³⁴ Smith never published these essays, but he exempted them from the arrangement he eventually made to have his papers burned upon his death.

primitive peoples who thought in such polytheistic terms—inhabiting a universe replete with gods, daemons, fairies, witches, and so on—did *not* perceive such entities acting to shape the “regular” phenomena of nature (e.g., the burning of fire and the falling of heavy bodies). Such events were part of “the ordinary course of things” that “went on of its own accord”³⁵ (*History of Astronomy* 48-50).

In the short essay on the history of ancient physics, Smith likewise faults the superstitious primitives for positing “designing, though invisible” beings to explain “almost every unexpected event.” As society progressed, fortunately, philosophy/science offered a superior vision (Smith equates philosophy and science), depicting the universe as “a complete machine...a coherent system, governed by general laws” that promote general ends: the preservation and prosperity of the system itself along with its various “species.” Such a universe resembles the machines that human beings produce, and philosophers (e.g., Timaeus and Plato) introduced “the idea of a universal mind, of a God of all, who originally formed the whole, and who governs the whole by general laws, directed to the conservation and prosperity of the whole, without regard to that of any private individual” (*History of Ancient Physics* 112-14). By positing a God that created an orderly universe whose laws are friendly to “species,” this theistic framework resembles the theology of *Moral Sentiments*, including Smith’s Providential account of the invisible hand and his frequent appeals to nature’s Author, Architect, Director, or Superintendent.³⁶

Smith in *Wealth of Nations* nevertheless evicts God, however tempted he might have been to argue along the following lines: *human* rulers must avoid deploying the visible hand of the state too aggressively since there is a *divine* wisdom that “superintends” the universe and promotes the “interest”

³⁵ For Smith, “nature” seems to mean the way something operates “of its own accord” (*WN* 372, 458, 523), without the intrusion of human violence, plan, constraint, artifice, or custom (28-9, 248, 265, 372, 489, 870).

³⁶ *TMS* 77, 93, 105, 128, 166, 169, 236, 289, 292. Smith links each of the three invisible hands to a broad pattern of socioeconomic development. Like *TMS*, the philosophical essays rely on a general contrast between savage and civilized society rather than on the four-stages theory of *WN*. The “notions” of the weak and fearful savage are “guided altogether by wild nature and passion” (*History of Astronomy* 49); philosophy/science only emerges in civilized society, where “law has established order and security, and subsistence ceases to be precarious”; “cheerfulness” and the consciousness of strength/security counteract the superstitious impulse to imagine “invisible beings”; with greater leisure, individuals who are “disengaged from the ordinary affairs of life” can be particularly observant (50); and opulence allows for the “evident distinction of ranks” that tames “confusion and misrule”(51).

of groups (especially nations) despite the selfishness and other shortcomings of so many individuals. The non-human authority/standard that Smith does retain is nature, as manifested in his pitch for the “natural system of perfect liberty and justice” (*WN* 606) that would support the “natural progress of opulence,” the “natural course of things,” the “natural progress of things toward improvement,” the “natural law of succession,” the “natural progress of law and government,” the “natural effort of every individual to better his own condition,” the “natural employments” of industry and capital, the “natural division and distribution of labour,” and so on. As sketched above, Smith insists that no “human wisdom” could equip “the sovereign” to superintend the industry of private people (687-8). Perhaps Smith abandoned theism in *Wealth of Nations* partly because of the threat posed by human rulers who restrict liberty while claiming access to some sort of divine wisdom.

Taken in isolation, however, the invisible hand of *Wealth of Nations* suggests that Smith remained willing to appeal to a non-human intelligence that superintends the welfare of at least human “wholes” such as societies (recall the philosopher’s God that secures “the conservation and prosperity of the whole”). Smith in *Moral Sentiments* repeatedly invokes “the wisdom of nature,” a phrase that highlights both nature’s intelligence and its capacity for “designing” (recall the distinction between the arbitrary “designing power” that superstitious people project onto gods and the philosopher’s God that “formed” and “governs the whole”).³⁷ But Smith mentions the wisdom of nature only once in *Wealth of Nations*. When criticizing Physiocrats who overestimate the importance of an “exact regimen of perfect liberty and justice,” Smith likens the “political body” to the human body, which contains “some unknown principle of preservation” that can protect our health against flawed regimens; the “wisdom of nature” can thus counteract “the folly and injustice of man” (*WN* 673-4). By linking nature’s wisdom to the “principle of preservation” that protects *bodies* (animal as well as human, one may infer), Smith signals another departure

³⁷ For the 1790 edition of *TMS*, Smith added a passage that evokes the spirit of *WN*. Consistent with the spirit of *TMS*, however, this passage still elevates a creator (above nature) who thinks, judges, arranges, and directs in order to promote the welfare of the whole: the “wisdom which contrived the system of human affections, as well as that of every other part of nature, seems to have judged that the interest of the great society of mankind would be best promoted by directing the principal attention of each individual to that particular portion of it, which was most within the sphere both of his abilities and his understanding” (*TMS* 229).

from *Moral Sentiments*, where he presents a world that is friendlier to human happiness, virtue, nobility, wisdom, love, benevolence, and tranquility.

In Smith's two books, the invisible hand is *not* an entity that superstitious people imagine in trying to comprehend disorder and frightening events. Rather, Smith formulates the phrase to help his 18th-century (and beyond) readers see reassuring types of societal *order*. Contra Grampp, the invisible hand *does* represent something "so complex and so grand as the social order" (Grampp 446). That order is not only broader than the inducement to employ capital domestically (supporting national defense), it is broader than Hayekian spontaneous order. For Smith, the order within a system of natural liberty is but one realm of invisible-hand dialectics.

Wealth of Nations innovates by depicting societal order in totally secular terms. But by invoking an invisible hand that *leads* people (he does not say that we are led "as if" by an invisible hand), Smith alludes to divine action. He thus invites attentive readers to focus on the book's treatment of religion, to notice the absence of God, and to contemplate the viability of both atheistic (*WN*) and theistic (*TMS*) worldviews.³⁸ Only *Moral Sentiments* attributes an Author to nature, and some of the differences between the two books may signal that Smith has used "invisible" authorial skills to "lead" his readers, especially when he appeals to God or nature as *authorities*.³⁹

³⁸ One can also approach the religious clash between *WN* and *TMS* by recalling the elusive dialectic *TMS* presents (in its invisible-hand chapter) between the "philosophical" view that condemns wealth/greatness and the "complex" view that celebrates them. The complex view emerges when "our imagination" leads us to confuse the "real satisfaction" that wealth/ riches and greatness/power provide with "the order, the regular and harmonious movement of the system, the machine or oeconomy" by which that satisfaction is produced (*TMS* 183). By highlighting our proclivity to become intoxicated by machines, Smith's "philosophical" critique of wealth and greatness might even prompt us to question the theism Smith celebrates in his essays (and in *TMS*); the theistic philosophers, analogizing from the unity and order of the machines that human beings create, portrayed the universe as "a complete machine" (*History of Ancient Physics* 113-14). Let me suggest one more conundrum. Insofar as Smith equates machines with "systems" (*TMS* 183, *History of Astronomy* 66, *History of Ancient Physics* 113) his "philosophical" critique of trinkets also poses a challenge to his own endeavors in formulating "systems" of political economy, moral philosophy, and jurisprudence (*TMS* 233-34, 265, 313-14, 340-42; *WN* 233, 606, 679, 687, 768-69, 780-81, 794). On the other hand, intellectual systems that resemble machines would presumably excel in precision, cohesion, reliability, and efficacy. Given the high standards that Smith thus set for himself, finally, perhaps Grampp (and other scholars) should work harder before concluding that Smith was a sloppy thinker or writer.

³⁹ The invisible hand can also remind us that, like our primitive ancestors, we are still prone to attribute agency to non-human powers that render us perplexed and puny.

Smith's essay on "the principles which lead and direct philosophical enquiries" incorporates rhetoric into its definition of science/philosophy. In a section that introduces his lengthy assessment of astronomy, Smith states that philosophy is "the science of the connecting principles of nature"; "by representing the invisible chains which bind together" the disjointed objects and events we encounter, philosophy tries to introduce order into the "chaos of jarring and discordant appearances," to restore the mind to "tranquillity and composure" (*History of Astronomy* 45-6). Just as some readers of *Wealth of Nations* doubt the existence of an invisible hand that leads people to promote beneficial ends, some readers of the astronomy and physics essays may be led to doubt whether human beings can attain *knowledge* of invisible *chains* that allegedly unify the cosmos. Smith proceeds to describe the historical essays in the following terms: "Let us" examine the different philosophical systems "without regarding...their agreement or inconsistency with truth and reality." Smith will merely assess "how far each of them was fitted to sooth the imagination, and to render the theater of nature a more coherent, and therefore a more magnificent spectacle." This rhetorical dimension, he adds, is what determines whether the authors "succeeded in gaining reputation and renown"; no system could attain "general credit" unless its "connecting principles" were "familiar to all mankind" (46).⁴⁰ After the long history of astronomy that culminates in effusive praise for the system of Isaac Newton, Smith concludes by apologizing, somewhat histrionically, for having ever implied that the "connecting principles" Newton presented were "the real chains which Nature makes use of to bind together her several operations" (105).⁴¹

As a reformer confronting a variety of powerful prejudices and interests that would inspire opposition to the new system of political economy he offers to the world, Smith might have felt compelled to employ exaggeration, irony, and other tools of persuasion: "If the rod be bent too much one way, says the proverb, in order to make it straight you must bend it as much the other" (*WN* 664). If "[a] philosopher is company

⁴⁰ The astronomy essay includes another remark that one can apply to the invisible hands of *WN* and *TMS*: in approaching a "strange" subject, Smith says, a writer could draw an analogy from a "familiar" one, creating not just "a few ingenious similitudes" but a "great hinge upon which every thing turned" (47).

⁴¹ If Smith in the 1750s was hesitant to claim that Newton had revealed the real but invisible *chains* that would "bind together" the movements of the planets, did Smith in 1776 believe that he himself had revealed a real but invisible *hand* that "led" lords, merchants, and others unwittingly to advance the "interest of the society," the "multiplication of the species" (*TMS* 185), "the publick interest" (*WN* 456) and the wealth of nations? In any case, Smith has left his readers with the additional challenge of reconciling natural *chains* with natural *liberty*.

to a philosopher only” (TMS 34), a philosopher’s books won’t always broadcast all of the complexities and uncertainties that fill that philosopher’s mind. In the 1790 edition of *Moral Sentiments*, Smith added praise of “the great wisdom of Socrates” (TMS 251), the philosopher who remains renowned for identifying his wisdom with his *ignorance* concerning “the greatest things” and for proclaiming that “the unexamined life is not worth living” (Plato’s *Apology* 22d, 38a). In the *Astronomy* essay, furthermore, Smith emphasizes that human beings pursue philosophy “for its own sake,” and that it began from “wonder” rather than from “any expectation of advantage from its discoveries” (*History of Astronomy* 51). Grampp may be wise in claiming that Smith’s allegedly “obvious and simple system of natural liberty” (WN 687) is “neither simple nor systematic and is by no means meant for all markets” (Grampp 442). But Grampp simply fails to appreciate how Smith’s invocations of an invisible hand can lead a reader to *seek* wisdom—from God, nature, prophets, philosophers, or other sources. Centuries after Smith’s death, we are still struggling to fathom a two-word phrase that stands out in a thousand-page book.

REFERENCES

- Bitterman, Henry J.** 1940. Adam Smith’s Empiricism and the Law of Nature, Parts I and II. *Journal of Political Economy* 48: 487-520, 703-734.
- Brown, Vivienne.** 1994. *Adam Smith’s Discourse: Canonicity, Commerce, and Conscience*. London: Routledge.
- Grampp, William D.** 2000. What Did Adam Smith Mean by the Invisible Hand? *Journal of Political Economy* 108 (3): 441-465.
- Griswold, Charles L., Jr.** 1999. *Adam Smith and the Virtues of Enlightenment*. Cambridge: Cambridge University Press.
- Hayek, Friedrich A.** 1948. *Individualism and Economic Order*. Chicago: Henry Regnery.
- . 1959. *The Constitution of Liberty*. Chicago, University of Chicago Press.
- Minowitz, Peter.** 1989. Invisible Hand, Invisible Death: Adam Smith on War and Socioeconomic Development. *Journal of Political and Military Sociology* 17 (2): 305-315.

- _____. 1993. *Profits, Priests, and Princes: Adam Smith's Emancipation of Economics from Politics and Religion*. Stanford: Stanford University Press.
- Muller, Jerry Z.** 1993. *Adam Smith in His Time and Ours: Designing the Decent Society*. New York: The Free Press.
- Rothschild, Emma.** 2001. *Economic Sentiments: Adam Smith, Condorcet, and the Enlightenment*. Cambridge: Harvard University Press.
- Smith, Adam.** 1776 [1776]. *An Inquiry into the Nature and Causes of the Wealth of Nations*. 2 vols. Ed. R.H. Campbell, A.S. Skinner, and W.B. Todd. Oxford: Oxford University Press.
- _____. 1776 [1759]. *The Theory of Moral Sentiments*. Ed. A.L. Macfie and D.D. Raphael. Oxford: Oxford University Press.
- _____. 1780 [1795]. The Principles Which Lead and Direct Philosophical Enquiries; Illustrated by the History of Astronomy. In *Essays on Philosophical Subjects*, ed. W.P.D. Wightman and J.C. Bryce, 31-105. Oxford: Oxford University Press.
- _____. 1780 [1795]. The Principles Which Lead and Direct Philosophical Enquiries; Illustrated by the History of the Ancient Physics. In *Essays on Philosophical Subjects*, 106-117.
- Viner, Jacob.** 1958. *The Long View and the Short*. New York: The Free Press.

ABOUT THE AUTHOR



Peter Minowitz is associate professor of political science at Santa Clara University, where he teaches courses in the history of political philosophy and was a co-founder of the SCU environmental-studies program. He received his PhD in political science from Harvard University in 1988. In addition to scholarly articles about Niccolo Machiavelli, Adam Smith, Karl Marx, Leo Strauss, Frank Herbert, Harvey Mansfield, and Woody Allen, he has published a book, *Profits, Priests, and Princes: Adam Smith's Emancipation of Economics From Politics and Religion*, with Stanford University Press. He is currently writing a book about the use and abuse of "diversity" in Catholic higher education, and he moonlights as a jazz pianist. His e-mail address is pminowitz@scu.edu.

“Credibility” in Context: Do Central Bankers and Economists Interpret the Term Differently?

JAMES FORDER*

A COMMENT ON: ALAN BLINDER. 2000. CENTRAL-BANK CREDIBILITY: WHY DO WE CARE? HOW DO WE BUILD IT? *AMERICAN ECONOMIC REVIEW* 90(5): 1421-1431.

[Abstract, Keywords, JEL Codes](#)

FOR MANY YEARS ALAN BLINDER HAS BEEN ONE OF THE outstanding analysts of the American economy and policymaking. In the 1990s he served as a member of the President's Council of Economic Advisers and then as vice-chairman of the Governors of the Federal Reserve. He has published and spoken widely on monetary policy and often expressed some measure of dissent, at least from what is widely held to be the consensus position of academic macroeconomists. In particular he has offered some of the most effective critiques of the “theory of policy credibility,” and made his position on it perfectly clear. In Blinder (1997, 13) he said, “I firmly believe that this theoretical problem is a non-problem in the real world.” Much the same message was repeated in Blinder (1998).

More recently, however, Blinder has taken a different tack. In particular, in Blinder (2000) he conducted a survey of academics and central bankers, asking them various questions about ‘credibility.’ Introducing this study, he said,

* Balliol College, Oxford.

In a word, credibility matters *in theory*, and it is certainly believed to matter *in practice*—although empirical evidence on this point is hard to come by because credibility is not easy to measure. This paper seeks to shed light on two main issues: Why and how? Why is credibility so important to central bankers? And how can a central bank create or enhance credibility? (Blinder 2000, 1421)

The survey consisted of a number of questions as to whether ‘credibility’ is important; what purposes it serves; and how it can be established. In each case a question was framed so that the respondent could state his strength of agreement with it on a scale of 1 to 5. Blinder then calculated the average score of the academics and the policymakers, and compared them, expressing some surprise about differences he found between the two groups.

In this comment, I suggest that an explanation of these differences is to be found in the two groups of respondents interpreting the questions differently, and that this is traceable to their having different understandings of what is meant by ‘credibility.’ Blinder himself notes that this word is used in a variety of ways but he chooses not to seek to clarify the meaning for his prospective respondents, since, as he put it, they “might have other meanings in mind, and therefore recoil from mine” (Blinder 2000, 1422). I suggest that there must be serious doubt about the value of a survey that asks the importance and ways of acquiring a thing without specifying what that thing is, but more importantly there is a further danger. Without clarity as to exactly what is meant by ‘credibility,’ some of the answers to Blinder's survey might appear to give support to proposals which are in fact unrelated to what the respondents had in mind. Muddle over the meaning of a word thereby contributes to creating the impression of there being wider support than in fact there is for these policy proposals.

If this suggestion is correct Blinder's results would retain their interest, but it would show that they also need the most careful interpretation before any conclusion could be drawn.¹ In particular, the conclusions that some other authors have begun to put on them would be inappropriate.

¹ Comparisons with other problematic words in the economics literature, such as ‘significant’ and ‘cause’ may be apparent.

THE MANY MEANINGS OF 'CREDIBILITY'

In ordinary, everyday English, the word 'credible' has a number of related meanings. When a policy or a policymaker is described as 'credible' that can denote trustworthiness, or competence, or a wide range of other things. The *New Oxford Dictionary* for example gives as definitions both "the quality of being convincing or believable" and "the quality of being trusted and believed in." The quality of being believed *in* suggests clearly capability, good judgment, and related characteristics. It also might be that in policy statements, 'credibility' is substantially a matter of intelligibility and of the private sector appreciating the policymaker's objectives and their understanding of the way the world works.

When something or someone is described as 'credible' in any of these senses—of intelligibility, effectiveness, capability, competence—that would certainly be a description which conveyed approval. In so far as 'credibility' is understood in this ordinary English sense, it would be bizarre to favor 'incredible' or 'non-credible' over 'credible' monetary policy.

Among monetary economists, on the other hand, the term 'credibility' has a special meaning, originating from the usage of Robert Barro and David Gordon (1983a). They gave no actual definition of the word, but used it to characterise policy which the private sector believes will be carried out in the context of a particular reason that might lead them to believe that it will not.² As a result, 'credibility' has become associated with certain features of the formal models that now fill much of the monetary economics literature.

The Barro and Gordon model is based on a vertical aggregate supply curve with rational expectations, together with an assumption of labour market or other distortions which cause equilibrium unemployment to be above its optimal level. They also assume that the marginal costs of inflation are rising, but zero at price stability. A consequence is that from that starting point, small positive inflation surprises are desirable because they lower unemployment, thereby moving it towards the optimum. Crucially, however, they assume that at least in the normal run of policymaking, there is no mechanism by which the policymaker can commit to particular future policy actions. As they argue, a consequence of this is

² In a number of their usages, the policymaker's credibility is specifically made to be a semantic equivalent of its 'reputation'—for example, Barro and Gordon (1983, 108) 'loss of reputation or credibility.' and "the government retains its reputation (or credibility)" (109).

that—again in the normal run of things—the private sector will not rationally expect the policymaker to deliver price stability since it would never be socially optimal for policy to conform to such an expectation. This gives rise to what is sometimes, following Kydland and Prescott (1977), known as the “time-consistency problem” which is that with expected inflation always positive, the policymaker can do no better than accommodate this expectation with the result that inflation is excessive, but there is no reduction in unemployment. The policymaker is, in effect, in the position of wanting to promise the private sector that prices will be stable but finding that such promises are, in the rational expectations context in question, ineffective in affecting expectations and therefore valueless.

The effect of this argument, as Barro and Gordon realised, is to focus attention on what can be done outside the normal run of things in order to induce the private sector to believe that policy will be set to achieve price stability. If the private sector can be made to believe this, policy will be improved because, although unemployment will remain above its optimal level, inflation will not. It is in this context that they used the word ‘credible’ to describe the characteristics of policy leading private agents to believe policymaker announcements about future inflation. In particular, they suggested that if one considers a “repeated game” version of the basic problem, reputational equilibria may exist with low inflation. Consequently they said that to the extent that an appropriate reputation can be established, ‘credibility’ for the low inflation policy is achieved. Later authors have, in proposing different solutions to the same problem, also described as ‘credible’ the policy, policymakers, or arrangements for making policy which are (in the models in question) successful in leading the private sector to expect price stability. Thus, in this usage a central bank’s or policymaker’s ‘credibility’ is the extent to which the private sector believe there will be low inflation.

One can clearly regard Barro and Gordon’s sense of the word ‘credible’ as falling within normal English usage, since if the policymaker actually promises price stability, one might, in ordinary English, say that if such a promise is credible, rational agents believe it will be honoured. On the other hand, their usage picks out only a small part of the range of meanings of the word.

First, Barro and Gordon’s usage does not in any way suggest general capability, reliability, or effectiveness: ‘credibility’ in their model is not a problem of policymaker being confused, unintelligible, or incapable; it is not achieved by those who are merely good at their job. Indeed, it is an assumption of their model that the policymakers they are considering are

well-intentioned, fully competent and wise to circumstances,³ and the point is to show that they nevertheless face a ‘credibility’ problem. So in Barro and Gordon’s usage ‘credibility’ picks out only a characteristic something like ‘honesty’ or ‘promise-keeping.’

But more than this, and most importantly, their idea of ‘credibility’ relates to promise-keeping only in the context of the particular incentive to deception that lies at the heart of the time-consistency problem. The issue is simply one of whether the private sector believes that the one particular incentive to create inflation described as the “time consistency problem” will be resisted. Policymakers who are believed to be likely to inflate specifically for the reason that it is in the public interest that surprises be created lack credibility. Those who are expected to resist this temptation have it. Therefore, other possible sources of confusion, dissembling, error, fudging, and other circumstances resulting in the breaking of promises—which are, unfortunately, common parts of policymaking, arising from, for example, the concealment of errors, the advancing of certain interests above others, and the achievement of electoral advantage generally—are not things under consideration anywhere in Barro and Gordon’s discussion of ‘credibility.’⁴ But any and all of these considerations might be unendingly at issue in ordinary-English ‘credibility.’

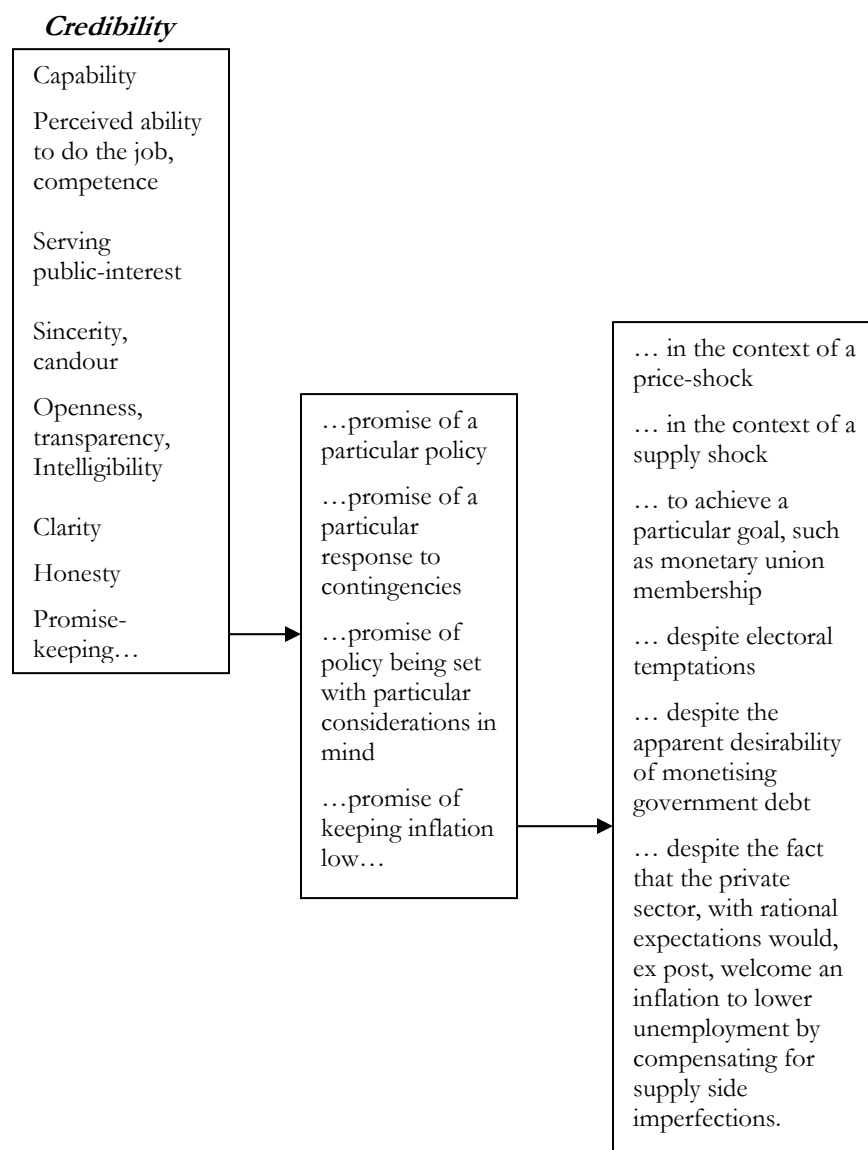
Figure 1 is an attempt to convey the many meanings that ‘credibility’ might have, and to show that the specific usage of that word in Barro and Gordon’s analysis—in the bottom right corner of the figure—is a very special one, and may not be what all the respondents had in mind in discussing ‘credibility.’ It can be seen first that there are many senses of “a credible policymaker” which indicate that the policymaker is in one way or another doing a good job. *One* of these is that it is thought that when the policymaker makes a promise, it can be relied on. But there are many promises one might make, and only one of these is a promise to keep inflation low. And furthermore, there are many reasons why one might wish to make that promise. Barro and Gordon’s discussion points to only one

³ One could go further and say the model would lose its point without this characteristic: there is no difficulty in explaining excess inflation if policymakers are presumed not to be competent. The assumption that they “act rationally” is made specifically in Barro and Gordon (1983b, 590), on which Barro and Gordon (1983a) builds, and is implicit throughout the latter paper.

⁴ Again, it would spoil the point of the argument if these things were under consideration. The power of the argument that equilibrium inflation is excessive arises from the fact that this is not attributable to such misfeasance. This point was made by Kydland and Prescott (1977).

such motive. That is that in their model there is *one particular* incentive to inflate, and the ‘credible’ policymaker is one whose promises not to give in to *that* temptation are believed. But that is a very special usage of the word.

Figure 1: Many Meanings of ‘Credibility’



BLINDER'S OBSERVED LACK OF TEMPTATION

The possibilities—and dangers—of confusion over the meaning of the word ‘credibility’ can be illustrated by considering Blinder's well-known claim that, contrary to what the Barro and Gordon argument would lead one to expect, when he was a policymaker he never “witnessed nor experienced” a ‘temptation’ to generate surprise inflation (Blinder 1997, 13). In the abstract there seem to be five reasonable explanations of never witnessing such a thing. They are:

1. Central bankers do not believe even unanticipated monetary expansion can lower unemployment.
2. They do not believe that equilibrium unemployment is socially sub-optimal
3. They do believe equilibrium unemployment is excessive, but do not want to lower it, even if there were no costs of doing so.
4. They believe surprise inflation would destabilize a repeated-game equilibrium like the one Barro and Gordon analyse, thereby destroy credibility in their technical sense, and thereby raise expected inflation and reduce the net present welfare-value of future outcomes
5. Along with inflation and unemployment, central bankers have another argument, of overwhelming weight, in their utility functions, namely, the avoidance of the breach of trust that deliberately inflating would constitute.⁵

Any of the first three possibilities might well bear examination, but none of them seem to point to the significance of ‘credibility’ on any likely definition. For current purposes, then, I take it that the crucial issue is between the fourth and fifth. In either of these cases, the policymaker might be appropriately described as ‘credible,’ but the questions of why credibility is important, and how one might achieve it would have quite different answers.

If explanation four is correct, one could say that the credibility problem in the Barro and Gordon sense exists, but the actual mechanisms

⁵ The avoidance of the appearance of incompetence could also serve as explanation (5), if a surprise inflation, though successful, were nonetheless deemed a case of incompetence.

for solving it are effective and taken for granted—the reputational effects are strong enough to deter surprise inflation. In that case, the maintenance and enhancement of reputation would clearly be a priority, and this might have implications for central bank appointments, the institutional status of the central bank, and its policy. The fifth case is crucially different. Here, even the one-period optimisation problem would be solved by a setting of monetary instruments consistent with stable prices.⁶ There would then be no credibility problem in the Barro and Gordon sense at all: no temptation to inflate arises, so there is no problem to solve, and no solution is needed. The implication is that the central banker would not feel impelled to do anything with a view specifically to establishing Barro and Gordon credibility. This view seems to be shared by Blinder himself.⁷ And the remarks of John Vickers (1998) point very much in the same direction. On the other hand, it is clear that this view does not in any way diminish the importance of finding central bankers who are manifestly competent, capable etc, and thereby sustaining ‘credibility’ in other important senses.

BLINDER'S SURVEY

Blinder's survey asked a variety of questions about the importance and sources of credibility, such as how important it is, in pursuit of which objectives it is important and by what means it can be acquired. I suggest one might well expect very different answers to many of these questions depending on whether the respondent treats the question as relating specifically to Barro and Gordon ‘credibility.’ I speculate further that central bankers have, like Blinder, never themselves felt nor witnessed among their peers a temptation to seek to exploit ‘surprise’ inflation, and that the reason is something like explanation five above. If that is correct, I suggest their natural response to questions about ‘credibility’ is to treat them, not as

⁶ There would remain an issue about the determination of private sector expectations of inflation, and hence unemployment. Even if the private sector do not know the policymaker's preferences, a long enough period of price stability will presumably lead to unemployment returning to its equilibrium level.

⁷ See for example Blinder (2000, 1427) where he says a central bank has a duty to be truthful, or Blinder (1997, 14) where he says that a sufficient solution to the time-consistency problem is to direct the central bankers not to aim at unemployment below the natural level, and that since they are responsible people, they will do as they are told.

questions about a problem they do not face, but rather as questions about ‘credibility’ broadly understood, meaning—roughly—“perceived capability, competence, and general honesty.” Academic economists, on the other hand, generally lacking these experiences, but trained in formal models and the attendant vocabulary, tend more to treat the question as being about Barro and Gordon credibility.⁸ In that case, the answer that ‘credibility’ is unimportant—a very odd response to a question about ordinary English credibility—becomes reasonable, since it is reasonable to doubt the value of the model. And indeed we see in Blinder (2000, 1422), not only did the academics rate the importance of ‘credibility’ lower than the central bankers, but their ratings had a much wider standard deviation: some people were dismissive of its importance.⁹

Blinder's question 2 asked the relationship between credibility and “dedication to price stability.” Those answering a question about capability and effectiveness may feel that the control of inflation is its key measure, and they will think the relationship is very close. On the other hand, those answering about the Barro and Gordon sense of the word face a dilemma. For them, credibility is associated with the private sector's *perception* of the central bank's inflation aversity; therefore it is reasonable to answer that there is only limited association with actual inflation aversity.¹⁰ This would again suggest that central bankers would on average give higher scores, which was the result of the survey.¹¹

Questions 3 to 9 were about *why* credibility is important. The outstanding features of the answers in this section of the survey are that central bankers' scores were higher for every suggested reason, and their average scores were tightly packed in the range 4 to 4.39. Thus they think credibility is important for every reason suggested in the survey. That is certainly consistent with the idea that they are tending to answer about ‘credibility’ in roughly the sense of capability or competence. The economists' scores were lower, ranging from 3.19 to 4.17, which

⁸ It is obvious, I hope, and certainly sufficient, that I am only conjecturing tendencies, not a great dichotomy between the two groups.

⁹ On the 5-point scale, the average score given by central bankers was 4.83, and the lowest was 4. The economists gave an average of only 4.23 with a standard deviation of .85 as against .37 for the central bankers.

¹⁰ The distinction is important in variations of Barro and Gordon's theme, such as, in particular, Backus and Driffill (1985). They consider the case of policymakers of various, but unobservable degrees of aversion to inflation and show that even a policymaker who is not particularly averse to inflation can sometimes establish reputation and thereby enjoy credibility.

¹¹ The average amongst the central bankers was 4.1 but amongst the economists only 3.31.

presumably again reflects the fact that they were thinking of the Barro and Gordon sense of credibility and some of them think it altogether unimportant.

Looking at the answers to these questions in more detail, with the alternative meanings of ‘credibility’ in mind, three otherwise puzzling results fall into place. First, Blinder notes the high standard deviation of the academics’ response to the idea that credibility is important to achieve disinflation at low cost in increased unemployment—the “credibility hypothesis” as he calls it—and second that the central bankers were much (and significantly) more impressed by the idea that credibility helps to keep inflation low than they were by the credibility hypothesis.

Both of these results should be seen in the light of the credibility hypothesis being a clear theoretical implication of the models of ‘credibility’ in the Barro and Gordon sense. In the models, wages are set on the basis of a rational expectation of inflation. Where credibility is high, even when inflation occurs (due to exogenous factors, presumably), the public believe that the policymaker is committed to reducing inflation, and set wages on that basis. Inflation is then reduced without an increase in unemployment. On the other hand, where credibility is low, higher wages are set, and if inflation is reduced, this occurs only at the expense of raising unemployment. Therefore, in theoretical approaches like that of Barro and Gordon, high credibility is associated with small increases in unemployment during disinflations. Those who are impressed by the theory will presumably believe that this is an important reason to value ‘credibility.’ On the other hand, economists who are sceptical of the value of the Barro and Gordon analysis are presumably more conscious of the empirical work cited by Blinder contradicting the credibility hypothesis.¹² Indeed, that work may be the source of their scepticism. Hence again, with some respondents thinking along the lines of the models, and others rejecting them, there is wide divergence of views on this point, generating the large standard deviation in economists’ responses.

On the other hand, those thinking of credibility in an ordinary way and particularly in a sense of meaning something like perceived competence, will certainly have reason to think that it is primarily valuable for avoiding outbreaks of inflation. Indeed, in many cases, allowing outbreaks of inflation might well be regarded as paradigmatic of monetary

¹² To which I would add Posen (1998).

policy incompetence.¹³ Hence, central bankers rate this as a highly important aspect of credibility.

The third puzzle was that the largest difference in scores between the two groups arose from the idea that “credibility is important as a way to justify public support for an independent central bank.” Economists gave this a rather low score and Blinder suggests it is because either they think independence is unimportant or that the public can be expected to support the independence of a non-credible central bank (Blinder 2000, 1427). Here I think he may have missed a third possibility. If we are using the word in the Barro and Gordon sense, one might say that the public are unlikely to understand what the ‘justification’ means, and therefore it cannot be effective in securing support. But in the ordinary English sense, credibility must be conducive to public support of independence—one can hardly imagine public support for independence being enhanced by incompetence and misfeasance. So it is no surprise to find the central bankers thinking credibility important to public support of independence.

The final part of the survey concerns the creation of credibility and asked a succession of questions as to how important certain things—such as independence or a record of honesty—are to the establishing of credibility. Here the outstanding result is that, taking the average of their scores, the economists and central bankers ranked the suggested means of acquiring credibility identically. That may not be a surprise even if the groups have a tendency to treat the question as asking different things. Blinder suggests that there is a puzzle here over the cases of the possibility of creating credibility by “living up to one’s word” or by central bank independence. Both groups rated the former higher than the latter, whereas Blinder suggest that the economics literature places so much emphasis on central bank independence that one might have expected it to be at the top of the economists’ list (Blinder 2000, 1427-8).

Some of the literature does indeed say that legislation creates credibility more or less directly, for example simply by granting independence to a central bank. But on the other hand much of it says that an independent central bank faces the same Barro and Gordon problem, so that ultimately only reputational solutions can be effective.¹⁴ Some of the literature also alleges a link between central bank independence and low inflation, but that is consistent with (amongst other things) either the view

¹³ There might be exceptions, of course, but it is unlikely to be said that every case of excessive inflation is adequately explained by external factors.

¹⁴ Lohmann (1998) argues this specifically, many others imply it in one way or another.

that independence directly creates credibility leading to price stability, or that independence creates the opportunity to build reputation. The academics' ranking here may therefore simply reflect a certain interpretation of the message of the literature: they may think that reputational, not statutory mechanisms are the ones which are effective in building "Barro and Gordon credibility." Indeed, this view would seem to be supported by the academics' low scoring of the ideas that credibility can be built either by adopting a policy rule or through creating mechanisms for central bankers to suffer personal loss when inflation is high (Blinder's questions 15 and 16). In either case, the proposal would be to create credibility by a legislative, rather than a reputational solution.

CONCLUSION

If my speculation—that different groups had a tendency to read the questions differently—is correct, Blinder's survey would be no less interesting. But it would create a danger if its results are not interpreted with the utmost care. In particular the survey risks giving the impression that there are kinds of uniformity of thinking between academics and practitioners which in fact do not exist. One effect of this might be to suggest that there is wider support than in fact there is for policy and institutional proposals designed to enhance 'credibility' in the Barro and Gordon sense. It would be a gross misinterpretation to suppose that this is what is being supported by those who believe central banks ought to be open, honest, and intelligible about their intentions and understandings, and that it ought to be manifest that policy is made dispassionately and skilfully. One cannot therefore welcome the argument of Issing et al. (2001, 37) referring to the working paper version of Blinder's paper. They say, "Central bankers are highly conscious of the benefits of credibility. Blinder's (1999) survey documents that they attach to this concept a higher importance than academics," and proceed by treating this as giving authority to their view that credibility in the Barro and Gordon sense is important. They then go on to treat this as giving support to proposals for firm commitments of policy designed to solve the particular problem Barro

and Gordon had in mind. Indeed, these are just the things that Blinder himself has said are a waste of time.¹⁵

REFERENCES

- Backus, D. and J. Driffill.** 1985. Rational Expectations and Policy Credibility Following a Change in Regime. *Review of Economic Studies* 52: 211-221.
- Barro, R. and D. Gordon.** 1983a. Rules, Discretion and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics* 12: 101-121.
- Barro, R. and D. Gordon.** 1983b. A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economy* 91(4): 589-610.
- Blinder, A.** 1997. What Central Bankers Can Learn from Academics—and Vice Versa. *Journal of Economic Perspectives* 11(2): 3-19.
- Blinder, A.** 1998. *Central Banking in Theory and Practice*. Cambridge, MA: MIT Press.
- Blinder, A.** 1999. Central-Bank Credibility: Why Do We Care? How Do We Build It? NBER Working Paper No. 7161.
- Blinder, A.** 2000. Central-bank credibility: Why Do We Care? How Do We Build It? *American Economic Review* 90(5): 1421-1431.
- Issing, O., V. Gaspar, I. Angeloni and O. Tristani.** 2001. *Monetary Policy in the Euro Area*. Cambridge: Cambridge University Press.
- Kydland, F. and E. Prescott.** 1977. Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy* 85: 473-491.
- Lohmann, S.** 1998. Reputational Versus Institutional Solutions to the Time-Consistency Problem in Monetary Policy. In *Positive Political Economy: Theory and Evidence*, ed. S. Eijffinger and H. Huizinga. Cambridge: Cambridge University Press, 9-22.
- Vickers, J.** 1998. Inflation Targeting in Practice: The UK Experience. *Bank of England Quarterly Bulletin* 38(4): 368-376.

¹⁵ See Blinder (1997).

ABOUT THE AUTHOR



James Forder is Fellow and Tutor in Economics at Balliol College, Oxford, Lecturer in Politics at Oxford University, a Managing Editor of *Oxford Economic Papers*, and a member of the Bar of England and Wales. He has written a number of papers critical of the theory of credibility and common interpretations of it, such as “The theory of credibility: confusions, limitations, and dangers,” in *Neoliberal Economic Policy*, ed P. Arestis and M. Sawyer. He has also been critical of the theory and practice of central bank independence, for example in “Some Methodological Issues in the Statutory Characterisation of Central Banks” in *West European Politics* 24(1): 202-216; “Interests and ‘Independence’: the European Central Bank and the Theory of Bureaucracy” in *International Review of Applied Economics* 16(1): 51-69; and “The Limits of ‘Independence’ and the Policy of the ECB” in *Public Choice*: forthcoming.

His email address is james.forder@balliol.ox.ac.uk

Identity and Politics in School Reform Research

FABIO ROJAS*

A COMMENT ON: GEORGE A. AKERLOF, AND RACHEL E. KRANTON.
2002. IDENTITY AND SCHOOLING: SOME LESSONS FOR THE ECONOMICS
OF EDUCATION. *JOURNAL OF ECONOMIC LITERATURE* 40(4):1167–1201.

[Abstract, Keywords, JEL Codes](#)

GEORGE AKERLOF AND RACHEL KRANTON (2002) START WITH a question: Why do schools with similar resources produce such widely varying outcomes? Why is it that schools with similar class sizes, for example, produce highly varied levels of student learning? Quoting one of their own referees, Akerlof and Kranton note that “economists do not have good models for explaining why school resources do or do not (as often found) affect the returns to schooling” (Akerlof and Kranton 2002, 1167). Economists have produced evidence showing that resources do matter, but they have not shown *how* they matter.

Motivated by the sociological literature on schools, Akerlof and Kranton make the following argument about school resources and achievement. Student learning depends on interactions between students and schools. Students have an identity, such as “burnout” or “nerd.” A student will try to learn at school if such achievement enhances his or her identity. Citing James Coleman’s (1961) study of ten Illinois high schools, Akerlof and Kranton present a model in which students,

choose their social category, and they choose effort in
school. . . . They consider the match between their own

* Department of Sociology, Indiana University

characteristics and the ideal characteristics of jocks, burnouts, and nerds. . . . They consider the match between their own actions and the ideal behavior of their chosen categories. (Akerlof and Kranton 2002, 1168–1169)

Therefore, school administrators will improve student performance if they use school resources to cultivate academically oriented identities among students. When student backgrounds provide an identity inimical to academic achievement, Akerlof and Kranton suggest that “the aim of the curriculum and other aspects of the reform is to trump the effect of students’ backgrounds on school participation” (Akerlof and Kranton 2002, 1189). Economists can better understand the effects of school resources if they view learning within a process of identity formation, resource allocation, and social interaction.

The sociological literature confirms the common sense view that students often care less about their studies than about what their friends think. This literature induces Akerlof and Kranton to formulate a model of student effort. Akerlof and Kranton adopt the standard approach to modeling the market value of education: skills are valuable because of market wages (see discussion of “student utility function” on page 1172). Akerlof and Kranton’s innovation, which first appeared in Akerlof and Kranton (2000), is to model the effect of student identity on the utility of class-work effort. The utility of effort depends on the returns to human capital *and* on how participation in academic activities reinforces the student’s identity. An identity/social category has an “ideal” effort level. “Jocks,” for example, would probably think that they should spend some effort on academic activities but the rest of their time on athletics, while “nerds” think that all their time should be spent on academics. Deviations from the ideal effort level decrease the student’s utility. For a student, the total utility of class-work effort depends on both market wages and the way academic effort figures into her identity.

Upon this model of student effort, Akerlof and Kranton build an analysis at the *school* level. They assume that a school has an “ideal,” which roughly means image or academic standard. In the model, the school ideal is a level of effort that teachers expect students to show. For example, at urban magnate schools teachers expect students to dedicate nearly all their time to school work. Akerlof and Kranton note that a school’s ideal might conflict with student identity. Students whose identities do not resemble the ideal presented by the school will “burn out” and choose not to exert the effort needed to learn. The proportion of “burnouts” depends on the

school's ideal and its position within the distribution of student identities. Therefore, school reforms will be effective when reform discourages burnouts by offering an ideal compatible with the student population. Akerlof and Kranton go on to argue that episodes of school reform can be explained as attempts to minimize burnouts and maximize student skill acquisition.

I criticize Akerlof and Kranton's school reform analysis on two grounds. First, Akerlof and Kranton assume that schools seek to maximize student skills; that assumption is, at best, incomplete. Second, Akerlof and Kranton misread the evidence regarding school reform.

SCHOOLS AS SKILL MAXIMIZERS

Akerlof and Kranton assume that schools choose to maximize the skills acquired by students. Akerlof and Kranton do not justify this assumption, but many of their arguments depend on it. For example, they suggest that a school might offer multiple ideals as an appeal to a heterogeneous student population, which should increase the school's average skill level. Two ideals correspond, for example, to having college-bound and vocational tracks in the school. Two ideals might also describe a situation in which teachers expect quality academic work from some students, while allowing other students to submit substandard work. Multiple ideals are meant to appeal to students with diverse identities in order to minimize the number of students who reject the school and choose not to acquire skills offered by the school. Akerlof and Kranton argue that when student populations are heterogeneous, schools offer multiple ideals: "the school achieves higher skills by providing two categories. Two categories increases the number of students who identify with the school" (Akerlof and Kranton 2002, 1187). Akerlof and Kranton state that "the more students view themselves as different from the school, the more the school must reduce its ideal to engage students in the school and increase skills" (Akerlof and Kranton 2002, 1186).

The skill maximization assumption completely ignores the fact that schools, especially public schools, have audiences and political constituencies that limit school operation and organization. In real life, the choice of an ideal—if such a choice is even made—is not necessarily an attempt to maximize mean skill levels. Sociologists and educational researchers provide

ample evidence that curricula and teaching practices are influenced as much by politics as by attempts to maximize marketable student skills. Historical analyses show that school organization emerges from a complicated series of decisions that are both pedagogical and political. For example, the Prussian school system, a model for the American school system, was designed mainly to teach loyalty to the nation-state (Bendix 1968, 244–245). Horace Mann, leader of the American Common School movement, urged that schools be created so that future generations would be saved from “vice” and prepared for their future “civic duties” (Cremin 1957, 75–77). Studies of recent curricular reform also focus on the political contexts of school organization. They emphasize that curricular change often depends on the reformers’ appeal to a school district’s political culture (Binder 2002).

In a single paragraph on page 1171, Akerlof and Kranton cite one unpublished article that recognizes schools as dependent on democratic politics (namely, Kremer and Sarychev 2000), but they fail to incorporate political realities into their description of how administrators and teachers operate their schools. Only in the last paragraph of their article do Akerlof and Kranton mention that schools might have goals other than maximizing student skills. Akerlof and Kranton note that parochial schools have non-economic goals: “religious schools often eschew economic goals in favor of religious goals. In some cases, they view their primary mission as the separation of the saved from the damned” (Akerlof and Kranton 2002, 1198). Akerlof and Kranton allude to the accusation that voucher-supported charter schools were created to maintain all-white schools: “A similar desire for separation lay behind the voucher-supported private academies established (unconstitutionally) in the wake of *Brown vs. Board of Education*.”¹ Because public schools also have non-economic goals determined by school boards and other political entities, Akerlof and Kranton write that “school choice may be neither skill-increasing nor ideologically neutral” (Akerlof and Kranton 2002, 1199). Because all schools have non-economic goals, choice may not make a difference. Akerlof and Kranton’s conclusion ignores the

¹ This accusation is surprising given that a Department of Education survey shows that charter schools often have *more* minority and limited English proficiency students. Analysis of the Department’s 1997 survey of charter schools shows that the student population in charter schools is nearly identical to the ethnic composition of the states where charter schools are located, except for Native Americans, who rarely attend charter schools (Office of Educational Research and Improvement 1998:10, N=428, response rate = 87%). The same survey found that 60% of charter schools have minority populations resembling their districts and 35% of them have higher concentrations of non-white students (11).

fact that public and private schools have different kinds of non-economic goals. Private schools must answer to parents and the communities that sponsor them, while public schools must cater to voters, local school boards, and teacher's unions. Akerlof and Kranton allude to this on page 1189 but fail to incorporate the insight into their school choice model, which leads them to believe that the ability to choose between public and private schools probably makes little difference. There is much value in presenting a simplified model of school reform, but Akerlof and Kranton's omission of the political contexts of the school is highly misleading and supports dubious conclusions about the efficacy of school choice.

MISREADING THE RECORD

Akerlof and Kranton's incomplete school choice model leads them to misinterpret school reform. For example, they make much of the fact that American schools have become "shopping mall" high schools (Powell et al. 1985) with lax disciplinary policies, and where students are allowed to choose courses. According to Powell et al., shopping mall schools best serve gifted and disadvantaged students in need of specialized education. Average students are not well served because they require a coherent and organized curriculum.

Akerlof and Kranton argue that their school reform model explains the shift to the shopping mall school. They use their model to interpret change in a school studied by Gerald Grant in *The World We Created at Hamilton High* (1988). Offering course electives and lax discipline means shifting the school's academic and disciplinary standards. Akerlof and Kranton's model depends on a parameter s , a value representing the ideal characteristics of students associated with S , a school's social category (Akerlof and Kranton 2002, 1183). Akerlof and Kranton note that "the choice of s involves a trade-off: Increasing s increases skills directly, but reduces $[1-\beta]$, the number of students who identify with the school. The optimal s balances these effects" (Akerlof and Kranton 2002, 1186). Akerlof and Kranton's model predicts that when students are heterogeneous, a single school ideal will produce too many "burnouts." They interpret the permissive high school as a response to an increasingly diverse student population.

Comparison of the total skills for one category, s^* , and of two, s_H^* and s_M^* , shows that when social distinctions are large (t is large or σ is large), the school achieves higher skills by providing two categories. Two categories increases the number of students who identify with the school. When the social distinctions are small . . . students are more likely to identify with the school, and a school will maximize skills by eliminating choice and providing a single standard. The events at Hamilton High are thus consistent with the model. In the short run, the single initial value of s is fixed and there was a massive disruption. This disruption eased in the long run as the school moved from a single- s ideal to a double- s ideal with choice and tolerance. (Akerlof and Kranton 2002, 1187)

According to Akerlof and Kranton, Hamilton High became a “shopping mall” after the school was forced to racially integrate, which increases σ , a variable denoting the diversity of the student population. The σ parameter reflects “socioeconomic status, race, and ethnicity” (Akerlof and Kranton 2002, 1183). Once black students entered the school, they severely misbehaved. The school shut down and reopened with more lenient academic and discipline policies. Akerlof and Kranton interpret this as a shift from a single ideal to multiple ideals. According to them, students had revolted because the school’s ideal was too different from the identities of the demographically diverse integrated student body.

It is true that the shopping mall high school emerged after a demographic shift, as Akerlof and Kranton’s theory predicts. However, Akerlof and Kranton omit details that contradict the purported explanation. For example, nowhere in Grant’s case study does anyone ever assert that permissive standards were introduced to reduce burnouts or maximize student skills. Instead, Grant’s case study shows that the permissive high school was ushered in by a combination of non-demographic factors. One factor was declining belief in teacher authority, a general cultural change starting in the 1960s. According to one guidance counselor who worked at Hamilton, students questioned teachers in the following manner: “A kid could always turn around and say to you, ‘How many of the faculty have an alcohol problem? And yet . . . they continue to teach.’ . . . So who are we to say what’s right or wrong?” Gerald Grant’s comment captures the key

point: "Here was the collapse of adult authority as a standard for children. Not only did this counselor express no moral authority, she actively concurred in the notion that adults in general deserve no authority because some adults have a drinking problem" (65). The guidance counselor failed to enforce the moral order of the school.

The reason that student disruption led to school reform was not changing student demography, but the disintegration of the cultural consensus undergirding teacher authority. Gerald Grant describes the shopping mall high school as a negotiation between angry students and teachers unwilling to assert their authority, not as a response to new student identities. Gone were the days when teachers were unquestioned authorities in their classroom. Teachers overreacted to lawsuits and state regulations curtailing their powers, and they were confused about what they could do in the classroom. Surveying teachers, educational researcher Henry Luffler found that "Courts have increased the insecurity of teachers as they deal with the average discipline problems that take place within a school... We found a great deal of misunderstanding about what courts have said, and an overwhelming pattern of overestimating the extent to which courts have told teachers what they can and can't do in the classroom" (Grant 1988, 203). A more likely story is that lawsuits, parental intervention, and classroom regulation became excuses to not enforce discipline, which is one of the teacher's most important and demanding jobs.

Cultural change in the 1960s and moral breakdown provide a more plausible interpretation of student misbehavior and its relationship to school organization than Akerlof and Kranton's theory of student demography, burnout and skill maximization. For example, Akerlof and Kranton present an exchange in which a student challenges a teacher as evidence that black students reject Hamilton High's ideal. In the exchange, the student has stood up in class to grab a piece of paper for a quiz, and the teacher insists that the student sit. In the exchange, the student says, "Why you picking on me? You don't pick on the white kids who borrow a piece of paper" (1184). Akerlof and Kranton use such episodes as support for their model: "In terms of our model, the diversity, σ , of the students rose; most of the new students' characteristics were considerably below the school's ideal s . The model predicts exactly what happened: the number of burnouts and the disruption in the school increased" (1184). I argue that this incident does not show burnout; the student hasn't rejected the school's academic standards. Instead, the student challenges the teacher's application of these standards, which Akerlof and Kranton briefly

acknowledge (Akerlof and Kranton 2002, 1184). Inconsistent standards are the justification for questioning the teacher's authority, not a rejection of the school's ideal. The dispute arises from notions of fairness, not burnout and student dis-identification with the school.

Consider the following: In the late nineteenth and early twentieth centuries, American public schools were extremely heterogeneous: they were filled with immigrants from dozens of nations. If the Akerlof and Kranton model is right in that schools shift their academic standards to reduce burnouts and maximize skills in heterogeneous student bodies, then early twentieth century schools would have dropped their focus on assimilation and allowed immigrants and their children to develop their own courses of study. The opposite happened: the presence of a heterogeneous student population was seen as a reason to force students to speak a single language and adopt the school's pedagogical mission. Why? American educational leaders were determined to have schools produce English-speaking Americans (Cremin 1957; Tyack 1974; Hampel 1986). The school's mission was determined by American political culture, not by a skill-maximizing response to student demography. Akerlof and Kranton discuss attempts to Americanize immigrant children but fail to integrate their observations into their school reform model. For Akerlof and Kranton, schools trying to Americanize immigrant children show "that even the most caring teachers can unknowingly offend their student and convey that they are inferior" (1181). Instead, I argue that this episode demonstrates the intransigence of early American schools in the face of a diverse student population.

CONCLUDING REMARKS

Akerlof and Kranton's model ignores the ongoing political forces behind school reform; it presents an unrealistic explanation for why schools change and why students revolt against their school. The model that Akerlof and Kranton should present is one that treats schools the same way Akerlof and Kranton treat students. School personnel (like university professors!) have goals reflecting not only utilitarian values but also their political and social identities. Public schools belong to a political apparatus and are inherently political, a point that Akerlof and Kranton acknowledge but do not fully incorporate into their model.

Akerlof and Kranton are right in pointing out that school resources can be used to “trump” local cultures that are hostile to school achievement, but they do not ask when teachers will use resources for student learning. They should recognize more that teachers have goals that interact with the goal of maximizing student skills. As Akerlof and Kranton emphasize, some goals, such as religious indoctrination, might assist students in learning because students will more closely identify with the school. In contrast, pursuing job security by limiting evaluation of what students have learned might work against skill acquisition. The efficacy of school reform depends not only on the deployment of resources but also on the trade-offs created by the pursuit of multiple goals.

REFERENCES

- Akerlof, George A., and Rachel E. Kranton.** 2000. Economics and Identity. *Quarterly Journal of Economics* 105(3):715–753.
- _____. 2002. Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature* 40(4):1167–1201.
- Bendix, Reinhard.** 1968. The Extension of Citizenship to the Lower Classes. In *State and Society: A Reader in Comparative Political Sociology*. Ed. Reinhard Bendix et al. Boston: Little, Brown, 233–256.
- Binder, Amy.** 2002. *Contentious Curricula: Afrocentrism and Creationism in American Public Schools*. Princeton: Princeton University Press.
- Coleman, James.** 1961. *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*. New York: The Free Press.
- Cremin, Lawrence, ed.** 1957. *The Republic and the School: Horace Mann on the Education of Free Men*. New York: Teacher’s College Press.
- Grant, Gerald.** 1988. *The World We Created at Hamilton High*. Cambridge, MA: Harvard University Press.
- Hampel, Robert L.** 1986. *The Last Little Citadel: American High Schools Since 1940*. Boston: Houghton, Mifflin.
- Kremer, Michael, and Andrei Sarychev.** 2000. Why do Governments Operate Schools? Mimeo, M.I.T.

- Office of Educational Research and Improvement.** 1998. *A National Study of Charter Schools*. U.S. Department of Education.
Online: <http://www.ed.gov/PDFDocs/chartsch98sum.pdf> (cited: November 2004).
- Powell, Arthur, Eleanor Farrar, and David K. Cohen.** 1985. *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Boston: Houghton, Mifflin.
- Tyack, David.** 1974. *The One Best System*. Cambridge, MA: Harvard University Press.

ABOUT THE AUTHOR



Fabio Rojas is assistant professor of sociology at Indiana University. He received his Ph.D. in Sociology from the University of Chicago in 2003. His main research interest is organizational analysis and its intersections with political sociology. His dissertation studies the diffusion of Black Studies programs among universities as an outcome of the Black Student movement. He also works on computational models of social systems and mathematical sociology. He applies computational techniques to the study of epidemics, consensus formation, and the diffusion of innovations. His email address is: frojas@indiana.edu.

Fashion Cycles in Economics

PHILIP R. P. COELHO, DANIEL B. KLEIN,
AND JAMES E. MCCLURE*

A COMMENT ON: WOLFGANG PESENDORFER. 1995. DESIGN INNOVATIONS
AND FASHION CYCLES. *AMERICAN ECONOMIC REVIEW* 85(4): 771-792.

[Abstract](#), [Keywords](#), [JEL Codes](#)

THE *RAISON D'ÊTRE* FOR SCIENCE, AND ECONOMIC ANALYSIS IN particular, is to understand and explain observable phenomena. Yet, in an article in the *American Economic Review*, Wolfgang Pesendorfer (1995) develops a model of the design cycle of fashion goods that shows little regard for data, observational experience, history, intuition, or semantic integrity. These cornerstones of scholarship are superceded by the quest for a mathematical argument sustaining certain preconceptions.

Throughout the paper there is an unmistakable animus against fashion; in the abstract, Pesendorfer notes: "The paper gives conditions under which all consumers would be better off by banning the use of fashion" (Pesendorfer 1995, 771). A quote from Georg Simmel reinforces this negative view of fashion: "Judging from the ugly and repugnant things that are sometimes in vogue, it would seem as though fashion were desirous of exhibiting its power by getting us to adopt the most atrocious things for its sake alone" (Simmel [1904], 544; quoted in Pesendorfer 1995, 771). This sets the stage for constructing a model where fashion serves no purpose other than social differentiation.

Relying on Simmel, Pesendorfer introduces fashion as a way in which people distinguish themselves from others.

* Coelho and McClure: Department of Economics, Ball State University.
Klein: Department of Economics, Santa Clara University.

The purpose of fashion is to facilitate differentiation of “types” in the process of social interaction. The demand for new designs is derived from the desire of agents to interact with the “right” people. (Pesendorfer 1995, 772)

Here the similarity with Simmel ceases. Confusion results from Pesendorfer’s going back-and-forth between “fashion,” “garment,” and “design.” The only aspect of a fashion garment that is important in the model is its “look” or design. Unlike real-world garments, Pesendorfer’s “design” does not comfort, protect, warm, or beautify; neither does it generate prestige nor ostentation. It is best to think of a “design” as something like a ticket that lets the buyer enter into interaction with other ticket holders. Pesendorfer uses the terms “fashion,” “garment,” and “design” interchangeably because, in the model, they all mean simply the ticket to mix with other ticket holders.

The design is the basis for a never-ending compulsory matching process that establishes two-person groups. There are numerous problems with the model: (1) There is neither a definition, nor an elaboration of, the phenomena modeled; (2) alternative explanations of fashion cycles are ignored and even deliberately elided;¹ (3) the model of fashion as a signal is so abstract that it has no observational counterpart in reality; and (4) real time does not exist, and neither memory nor history exist.

¹ Pesendorfer draws heavily upon Simmel, but misquotes Simmel in a way that makes it appear that Simmel’s ideas are as one dimensional as Pesendorfer’s. On page 771, Pesendorfer shows a block quotation from Simmel. But looking up the quote, we found that he had edited out, without inserting ellipse to indicate omissions, “the desire for change” as an impetus for fashion. Below is the exact quote from Simmel ([1904], 543) with the text omitted without indication by Pesendorfer shown in bold italics:

[Fashion] satisfies ***in no less degree*** the need of differentiation, ***the tendency towards dissimilarity, the desire for change and contrast, on the one hand by a constant change of contents, which gives to the fashion of today an individual stamp as opposed to that of yesterday and to-morrow, on the other hand*** because fashions differ for different classes—the fashions of the upper stratum of society are never identical with those of the lower; in fact, they are abandoned by the former as soon as the latter prepares to appropriate them.

A DISCUSSION OF PESENDORFER'S MODEL OF FASHION DESIGN CYCLES

Pesendorfer's theory has two parts: 1) a static matching model that is constrained to two types of consumers; and 2) a dynamic game where a designer sets the prices of new designs and decides when to innovate in response to a dynamic version of the static matching model.

The Static Matching Model

1. *High is better than low*

The model assumes two types of consumers (high and low types), both preferring to be matched with those who are high types, but the high types cannot be identified by either observation or reputation. The "purpose of a consumer" in Pesendorfer's model is "to 'date' another consumer" (Pesendorfer 1995, 775). Mathematically the high types are in the range $[0, \alpha]$, and low types $[\alpha, 1]$.² "Type" is left undefined; it can be almost anything imaginable. "Depending on the interpretation, the type of an individual [consumer] may refer to her education, entertainment skills, or human capital" (775). The meaning of neither "date" nor "type" is specified. In each period all individuals are assigned partners, forming dyadic matches. It is unspecified who, or what, establishes and maintains the matches. Once matched, an individual is stuck with that partner for that period. Specific partners can neither be chosen nor rejected; partners are assigned dictatorially in a system of forced association.³

Both high and low-types prefer to be matched with a high-type. The utility premium associated with a match with a high type is greater for high types than it is for low types. In other words, what a high-type gains by avoiding a low type is greater than what a low-type loses by not being

² The functions that are used to depict consumer behavior are continuous. This means that there are an infinite number of consumers and/or the consumer is infinitely divisible into smaller units.

³ When Pesendorfer presents the matching rule, he says that individuals paired according to the rule "meet" (Pesendorfer 1995, 776) each other. In the abstract he refers to his theory as a "dating game" (771). In fact pairs of people not only meet each other but are assigned to be together for the duration. His usage is like suggesting that two prison inmates who are assigned to share a prison cell for the duration of their sentence merely meet each other. Pesendorfer's word choice obscures the assumption of forced association.

matched to a high-type.

The money price paid for a design is subtracted from the buyer's contemporaneous utility. A differential in utility between high and low types when matched to a high-type is necessary for the existence of equilibria where high-types pay for stratification and low-types do not pay.

Although Pesendorfer states that the matching model is applicable to the market for high-fashion garments (Pesendorfer 1995, 772), one of its fundamental assumptions turns reality on its head. Under Pesendorfer's assumptions, a high-type is willing to pay more to for a design the more that other high-types are wearing the identical design. Yet in reality, if a fashion-minded person goes to a social affair and finds others bedecked identically to herself, she is like to be disconcerted, rather than glad, about matching outfits.

2. *Buying designs and matching rules*

It is assumed that there is only one design that individuals may buy, and a single institution that issues it and sells it. Consumers desire the design because people who buy it are (randomly) matched with others who have bought it; people who do not buy the design are (randomly) matched with others who do not have it.⁴ If only high-type people have the design, then buying it ensures a high-type partner with certainty. If 80 percent of people with the design are high-type, then, irrespective of whether you are high or low, buying it provides a 0.8 probability of a match with a high-type partner.

If everyone but one buys the design, then the exception is matched to a low-type person. Pesendorfer does not address the implications of this assumption; there are several possibilities: perhaps there is cloning on demand for a low-type, or perhaps there is a freezer from which low-types are drawn and thawed as needed, or, more realistically, there could be a dungeon where low-type people are kept waiting for such exigencies. There are more difficulties. Who oversees the cloning/freezer/dungeon—and

⁴ In presenting the matching rule, Pesendorfer misstates the meaning of his notation, making the model difficult to understand. He uses $\mu_i(n)$ to mean the amount (i.e., measure) of type i consumers (i being high or low) using design n as a fraction of the set of *all* consumers (a set which has measure equaling 1). For example, if high types constitute a third of all consumers (that is, α equals $1/3$) and they all buy design n , then $\mu_h(n)$ equals $1/3$. Yet here is how Pesendorfer explains $\mu_i(n)$: "Let $\mu_i(n)$, $i = 1, h$ denote the fraction of consumers of type i using design n " (p. 776). What he says would imply for the example that $\mu_h(n)$ equals 1, rather than $1/3$.

why? The assumption of a slack low-type means that low-types always attach a positive price to the design. The design would have no value to a lone person without it if she were matched randomly with the set of all others; this would drive the price to zero. To prevent this, Pesendorfer has to assume that a *low-type* match will be forced (with certainty) upon the isolated design-less consumer, a match that (in equilibrium) will always make him worse off (in expected value) relative to parting with money and acquiring the design (giving him a chance of being matched with a high type). The model that Pesendorfer constructs is one where fashion is “useless” except in a “social context” (Pesendorfer 1995, 775). The assumption of forced association in Pesendorfer’s model drives the consumer demand for design. Producing more than one design and competition are “wasteful” (783).

In the parallel case, in which only one person buys the design (as opposed to all but one), then he is matched by the same rule that applies when no one has bought a design. Consequently, there is no need to hold a high-type person in reserve. Pesendorfer does not comment on the asymmetry of assumptions about what happens in the alternate cases where one person decides differently from all the rest.⁵

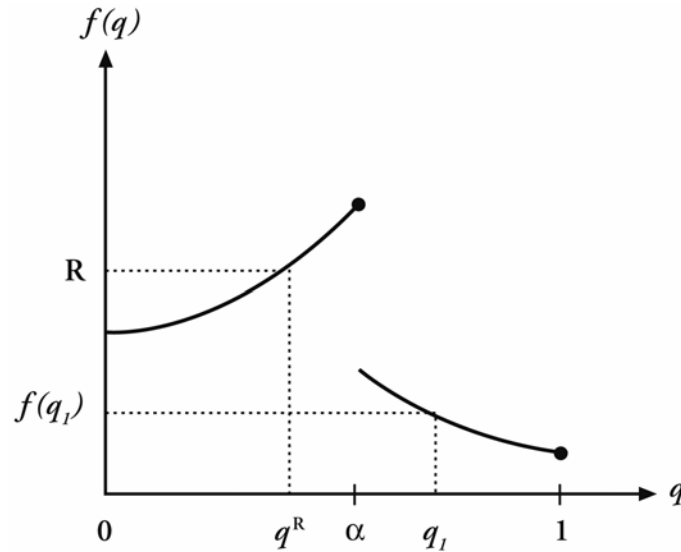
Before leaving the matching rule, there remains the issue of semantic integrity. Pesendorfer repeatedly calls the matching rule “the matching technology” (Pesendorfer 1995, 776-777). But normally “technology” means applied science or know-how. The matching rule in Pesendorfer’s model has nothing to do with know-how deployed to accomplish a goal; it is not technology, simply an assumption about the way things are. Calling the matching rule “the matching technology,” suggests that the fashion industry somehow uses technology to coordinate the matching of people, and obfuscates the artificial assumption of forced association.

3. *The last-to-buy function*

Pesendorfer’s Figure 1 (Pesendorfer 1995, 777), redrawn with embellishments here as Figure 1, elucidates the model; the horizontal axis measures consumers (denoted by q). Again, the high types and the low types are divided by α ; those on $[0, \alpha]$ are high, those on $(\alpha, 1]$ are low.

⁵ The symmetric assumption would be that when just one person buys a design, that person is (with certainty) matched with a high type, who is otherwise held in the dungeon.

Figure 1: The If-I-Were-Last-To-Buy Function



The vertical plot is $f(q)$: the expected benefit (willingness to pay) that consumer q would get from a design if she is the last one to buy a design (that is, if the design is bought by everyone on $[0, q]$ and no one on $(q, 1]$). For example, consumer q_1 would be willing to pay up to $f(q_1)$. In this model, *everyone's benefit* depends on the choice at the margin. Whether consumer q_1 (or some nonzero measure of consumers in the neighborhood of q_1) chooses to buy the design will affect the benefit of all other consumers. More specifically, when q_1 is a low type, it will reduce the benefit of all the interior consumers. Unlike the marginal benefit curve for apples, $f(q_1)$ does not display the benefits of *inframarginal* units. It shows one's benefit only when one is the "last" to buy it. If consumer q_1 is the last to buy, the actual benefits for all the interior consumers are as follows: (1) for the low type consumers with the design (that is, those on $(\alpha, q_1]$) the benefit is a *flat* line at $f(q_1)$; and (2) for the high types (on $[0, \alpha]$) the benefit is a *flat* line that is higher than $f(q_1)$. It is higher because high types benefit more from being matched with high types.

Notice that $f(q)$ is initially upward sloping. Consider the condition of consumer q^* (such that $0 < q^* < \alpha$), and assume that she is in the position of being the last one to buy the design. A high-type consumer will achieve the higher utility associated with matching with another high-type

consumer. But the consumer's willingness to pay for a design varies because the utility associated with not buying the design depends upon how many others have bought the good. Over the interval $[0, \alpha]$, as q^* approaches α , the pool of people left without the design (every $q > q^*$) is increasingly comprised of low types. As q^* goes up, the last high-type is buying it because if she does not, then the likelihood of a match with a low type rises.

Although we are calling $f(q)$ the last-to-buy function, Pesendorfer repeatedly calls it "the demand function" (Pesendorfer 1995, 776-777). The definition of a demand function is a function which, given a price, specifies the quantity demanded. In Figure 1 consider price R for the design; if $f(q)$ were a demand curve, that would mean that, at price R , consumers would demand q^R of the design. But, in fact, at a price of R , there are *three* equilibria: 1) an equilibrium in which no one bought the design (no one wants to be the only person to shell out for a design because in that case it does nothing to advance the quality of the forced match);^{6,7} 2) an equilibrium where every high-type (and no low-type) bought the design (in that case high-types would get benefit at the apex of the $f(q)$ curve and pay R for it); and 3) a knife-edge unstable equilibrium with exactly q^R consumers buying the design. When q^R consumers buy the design, the benefit exactly equals R , so in fact each high type is indifferent between buying and not buying.⁸

Pesendorfer does not explain the three equilibria. In calling $f(q)$ the demand function he suggests that at a price such as R it is the knife-edge unstable equilibrium that prevails. But Pesendorfer does not, in fact, treat $f(q)$ as a demand function. In all of his equilibria, if one high-type buys, they all do. Pesendorfer simply misuses the term "demand function," generating much confusion for the reader.⁹

⁶ Incidentally, it seems to us that, contrary to Pesendorfer's expression (4) on page 776, $f(q)$ equals zero when q is 0.

⁷ This equilibrium is unstable in the sense that if any non-zero measure of consumers "trembled" and purchased the design, dynamic forces would not push the system back to the original equilibrium. However, it is stable in the sense that it is not a knife-edge equilibrium. Every non-zero measure of consumers wishes to stick to the equilibrium not merely out of indifference, but of strictly superior utility. Thus, one might say that it is an unstable equilibrium, but not a knife-edge unstable equilibrium.

⁸ For the prices corresponding to the lower segment of the function, the "demand function" interpretation works—given a price, the function tells you quantity demanded, in a unique and stable equilibrium.

⁹ Pesendorfer makes another error when he writes: "This [the upward sloping portion of the so-called demand function] can be interpreted as a 'bandwagon effect' (Leibenstein, 1950)." Pesendorfer's attribution is flawed because Leibenstein explicitly modeled the bandwagon

Dynamic Equilibria and Design Cycles

To generate periodic design cycles, Pesendorfer embeds a modified version of the static matching model into a dynamic game in which a designer innovates and sets prices.¹⁰ Regarding the dynamics of consumers acquiring and processing information, in the appendix are the following remarks:

I assume that all agents can observe other's actions. However, strategies will be required to be anonymous (i.e., the deviations of a measure-zero set of agents do not affect equilibrium outcomes). Note that the interaction between consumers is entirely determined by the matching technology which determines matches using the currently displayed designs of consumers. Therefore, information about individual consumer histories is irrelevant for all agents, and one can interpret the game as one in which only the designer's action and total sales can be observed. (Pesendorfer 1995, 787)

The parallel discussion in the text is:

Implicit in the definition of the game is that histories of individuals are unobservable and hence the matching technology cannot condition on past designs used by a consumer. (Pesendorfer 1995, 778)

On one hand, there are zero costs of discovering what everyone else is currently consuming, but on the other hand there is neither memory about what was consumed in prior periods, nor strategic adaptation as new information is acquired. In this world “dating”/“matching”/“meeting” is

effect so as to exclude the existence of upward-sloping demand functions.

¹⁰ The mathematics of this game are so involved that even the paper's three page appendix is inadequate; readers who are interested in a complete explanation must acquire one of Pesendorfer's working papers. The working paper adds nuances that only lead to a greater number of possible equilibria. Because the nuances make the model even less operational, we do not examine it further. Incidentally, in condition (i) on page 779, we believe it should say $p \leq P(q)$, not $p \geq P(q)$.

behind a veil of never-ending ignorance, where it is impossible to learn from the past, hence the “technology” of human interaction is immutable. Here there is only one thing that can be observed about other people, the “look of,” or design of, a thing that everyone, without exception, will look upon as either “in,” or “out,” of fashion.

Discussing what happens in each period, Pesendorfer states that, “then [in each and every period] each consumer decides which designs to buy and which (of the designs she already owns) to sell” (Pesendorfer 1995, 777).

In footnote 10 Pesendorfer states: “All results are unchanged if consumers are not allowed to sell their designs.” Ordinarily we would expect a change in consumer behavior if the rights to resale are abrogated. But in Pesendorfer’s model: “If the new design is sold to both high and low types, then the old design *must* have a zero price” (Pesendorfer 1995, 778, emphasis added). The following paragraph explains why it must have a zero price.

If a consumer is the only person to purchase a design, then she will be matched with a random consumer from the pool of individuals who use no design. The design does not improve the quality of the consumer’s match in this case and therefore has no value. Thus, a design is only valuable to consumers if a coordination problem is solved. I assume that the designer can coordinate demand for his latest design. Part of the innovation cost c should be interpreted as expenses for marketing and advertising to achieve the coordination of consumers to the largest demand. I also assume that whenever the designer creates a new design he cannot simultaneously advertise the old designs, and hence the coordination of the demand for the old designs breaks down. Consequently, I restrict attention to equilibria in which designs other than the latest innovation are sold at a zero price. (Pesendorfer 1995, 778)

Here Pesendorfer means that: 1) People have no discretion over whom they are to associate with, instead they are randomly matched (noted previously); 2) A “coordination” problem has to be overcome, otherwise designs (which are assumed to contribute nothing to consumer utility directly), would be valueless to consumers; 3) The designer, via “advertising and marketing,” is able to “coordinate” consumers so that an otherwise

valueless design takes on value because the designer is given the power to force consumers to make the choices that give value to the design; 4) Although the designer's ability to coordinate is unlimited for the latest design (even in cases where the "latest" design lasts multiple periods), it is assumed that no one is able to coordinate the consumers of old designs to any degree (this forces the prices of old designs, in "equilibria," to zero).

Coordination costs are assumed to be: 1) a component of fixed cost incurred at the time of innovation; 2) invariant with respect to the size of the market and the proportions of consumer types. Although a number of coordinations may be required in multiple periods (in the cases where the "latest design" lasts more than one period), the costs of coordinating in all periods for a particular innovation occur at just the moment of innovation. This assumes that the designer has perfect foresight and that he prefers not to defer deferrable costs. The fixed cost "c" in Pesendorfer's model is a catch-all "black box" holding both the costs of design "innovation" and the costs of coordinating consumers. This is where the wastefulness comes in: something that is valueless to consumers ends up being innovated, and in some of Pesendorfer's cases, innovated an infinite number of times. Pesendorfer claims that "the current model predicts 'overinvestment' in product quality" (Pesendorfer 1995, 775), but really the result drops out directly from the bizarre assumptions about "matching" and the sterility of "designs." The result is not "predicted," it is deliberately constructed.

"STYLIZED FACTS"

Some of Pesendorfer's "stylized facts" (Pesendorfer 1995, 785) warrant special scrutiny; one is the crucial characteristic that drives the markets for durable fashion goods.

Appearance is an important component of most durable consumption goods. Large amounts of resources are devoted to the development of designs for clothing, cars, furniture, and electronic equipment. These resources are not primarily used to make those goods more functional; rather their goal is to let the product appear *fashionable*. (Pesendorfer 1995, 771, italics in the original)

Another “stylized fact” is that monopoly power exists in fashion designs.

[I]t is observed that even if potential competitors are free to enter the design market, *one* possible outcome is that one designer is chosen to be a fashion czar and behaves as a monopolist. If all consumers believe that only the fashion czar is capable of creating “fashion,” then this will be the equilibrium outcome. (Pesendorfer 1995, 773)

History and other types of evidence contradict both the reasonableness of focusing solely upon designs and the assumption of design monopolies. The model assumes that an innovation in fashion (design) has a temporal monopoly. Pesendorfer suggests that over time the monopoly breaks down as the design is copied and becomes less fashionable because it is more widespread. He specifies that the temporal monopoly is about a year long for clothing design (Pesendorfer 1995, 785).¹¹ These “stylized facts” are incompatible with the facts. If a new design appears in Milan or Paris, that design may be transmitted across the globe in minutes. Sketches, photos, specifications are all in the hands of competitors less than one day after an initial showing. Indeed, the technology of hand-held computers, cameras, and camcorders allows photos of designs to be transmitted worldwide within seconds.¹²

Nor did the stylized facts apply to the fashion world before the advent of real-time telecommunications: the resources invested in

¹¹ Pesendorfer (1995, 771) states: “A model of fashion cycles is developed in which designs are used as a signaling device in a ‘dating game.’ A monopolist periodically creates a new design. Over time the price of the design falls as it spreads across the population. Once sufficiently many customers own the design it is profitable to create a new design and thereby render the old design obsolete.”

¹²The Fax machine was widely available in the 1970s insuring the virtual instantaneous transmission of new designs. Given the importance of current design or fashion in the garment industry, substantial resources are committed to discovering what the new design's are before they are shown. Consequently some of the competitors have knowledge of new fashions even *before* they are officially shown for the first time. Some employers restrict their employees from bringing hand-held computer devices to work because of the ease with which designs can be transmitted.

discovering fashion trends ensured that this information was quickly and widely disseminated as far back as reliable information about a fashion industry exists. E.L. Jones, writing on the fashion markets of the eighteenth century, states:

In important respects Europe had become a unified market area, for the factors of production, capital and labor, and increasingly for goods. Obviously it was easiest for commodities of low bulk and high value to surmount the physical and political obstacles to trade, as we may see from the world of fashion. Unification in this field was aided by the Grand Tour and the practice of sending annually round each major city as far afield as St. Petersburg, and that western outpost of Europe, Boston, Massachusetts, a jointed wooden doll dressed in the season's Paris mode. Dressmakers everywhere copied the style. Generals permitted the 'wooden mademoiselle' to pass through the lines. Fashionable society was animated by periodic crazes for foreign styles, in the continent for things English, in England for things French. Scarcely a corner was remote enough to be exempt: even the National Museum of Iceland contains furniture which represents Danish taste, itself swayed by French, English, German and Dutch influences. (Jones 1981, 113-114; citations omitted)

If dressmakers in eighteenth century Boston were using that year's Parisian styles, how can a temporal monopoly in fashion design exist in the present-day?

The answer is that no quasi-monopoly in fashion design exists. What principally allows garment producers to price their products at a premium to "ordinary" garments is their reputation for producing superior garments, superior in a number of production characteristics. Some are qualities associated with: (1) the fabric, such as type (wool, cotton, linen, silk, blend, synthetics, etc.), weave, thread count, weight, color, backing, and so forth; (2) construction (double or single thread) and piping; and (3) ancillary objects (buttons, zippers, ornamentation). Whole sub-industries are devoted to, for examples, buttons and zippers, and an incorrect decision

on one of these margins can force the garment maker into closure.¹³ A major reason why fashion goods sell for premium prices is that they are relatively expensive to produce because the materials and specialized production capabilities that produce fashion goods can only be supplied at positive and usually increasing marginal costs. In contrast Pesendorfer's model requires that the marginal cost of an additional garment is zero.¹⁴

People in the fashion trade realize that the marginal cost of a particular design is trivial. Consequently consumers do not pay for design *per se*. Because Pesendorfer's model is about how consumers pay for fashion design, it is either devoid of empirical content, or easily refuted by looking at a recent issue of a fashion industry trade journal. Pesendorfer seems to acknowledge this in his concluding remarks. On the basis of the comments of an anonymous referee (Pesendorfer 1995, 786, n30), he states that the costs of "materials" applied to the designs are essential considerations to fashion theories that could be observationally tested (786-787). He leaves these considerations "for future research" (787).

To support his "stylized facts" and the explanatory power of his model, Pesendorfer refers to sumptuary laws of the past (Pesendorfer 1995, 786, n29). Pesendorfer argues that "sumptuary legislation will be efficient in the sense that the maximal gains from social interaction will be realized without waste of resources on design innovations" (786); thus the sumptuary laws are something he believes his model explains and justifies. There are some problems here. First, sumptuary laws did limit imitation of the "upper" class by the "lower" class, but sumptuary laws limited the lower classes' access to *products* (silk, gold thread, certain foods and colors, etc.). Designs, *per se*, were unimportant.¹⁵ Also, Pesendorfer says we would expect to find sumptuary laws "in societies with a well-defined class structure" (786), yet his assumption about individuals knowing nothing about each other except the design of their clothing especially lacks plausibility if class structure is well-defined.

¹³ One of the co-authors (Coelho) has had real world experience in this area. His firm sold fashion blue jeans, and he contracted to have cheap, yet purportedly high-quality zippers placed in the garments that were being manufactured to specifications. The zippers were cheap, but turned out to be low-quality; they had a tendency to split apart when worn. The subsequent purchase-returns and loss of goodwill led to the firm's closure.

¹⁴ "Once a design *n* has been created, the designer can produce indivisible unit of it at zero marginal cost" (Pesendorfer 1995, 775).

¹⁵ In his own discussion of sumptuary laws, Pesendorfer (1995, 786), citing J. M. Vincent (1934), alludes to the case of lower classes during the Middle Ages being restricted from wearing "velvets and silks", but nowhere does he provide an example of sumptuary laws on designs *per se*.

FINAL REMARKS

Pesendorfer's fashion model involves a long chain of complex mathematical relationships with vague linkages. Mathematical modeling can be useful in elucidating complex reasoning where verbal or descriptive reasoning is not sufficiently precise, but modeling complexity in economics comes at substantial cost. An underlying assumption of mathematical reasoning is that the relationships are stable throughout the analysis. The longer the chain of mathematical reasoning required by a theory, the longer the chain of required stable relationships between the variables. The difficulty with this stability assumption is that economic analysis deals with phenomena that occur in the real world in real time. Operationalism, the ability to assess models against real-world observations, is crucial in economic model building.

In 1955 Donald F. Gordon identified the problem with employing long, complex mathematical relationships in economic models. Each mathematical step is temporally stable only by virtue of the *ceteris paribus* assumption. The longer and more complex the mathematical model, the more likely the assumption of *ceteris paribus* will be incorrect. As long as the workings of the model do not occur instantaneously the passage of time materially affects outcomes in ways unspecified by the mathematics. Gordon noted a paradox: mathematics is most useful in elucidating long, complex chains of reasoning, yet the longer the chain of reasoning the more likely that *ceteris paribus* assumptions will be violated. Paradoxically, the more useful mathematics is in framing explanations, the more likely is it that observational reality will confound the explanations.¹⁶

Fashions are notoriously ephemeral. What can be said of the length of the mathematical chain in the paper? It has 10 numbered equations (excluding those in the appendix), one mathematically stated theorem, three mathematically stated propositions, and a mathematical appendix

¹⁶ Paul A. Samuelson (1952, 57) noted that Alfred Marshall and John Stuart Mill were given to "speaking of the dangers involved in long chains of logical reasoning." He explained Marshall's perspective: "Marshall treated such chains as if their truth content was subject to radioactive decay and leakage – at the end of n propositions only half the truth was left, at the end of a chain of $2n$ propositions, only half of half the truth remained, and so forth in a geometric multiplier series converging to zero truth." Gordon (1955, 58) also cited Marshall as having disdain for "long chains of reasoning." Also alert to the problem of employing tenuous derivations in economics, Wassily Leontief (1971, 1-2) remarked that: "Uncritical enthusiasm for mathematical formulation tends often to conceal the ephemeral substantive content of the argument behind the formidable front of algebraic signs."

(Pesendorfer 1995, 787-791) that proves the propositions, but only begins to prove Theorem 1. In the appendix Pesendorfer indicates that only the “outline” of the proof of Theorem 1 will be provided; on the same page, in footnote 32, he directs readers who are interested in the “details” of this proof to a 1993 working paper of his (Pesendorfer 1995, 788). The working paper is thirty-six pages in length including references and figures.

The costs of mathematical complexity in economics can be partially or entirely offset by the additional insights that the mathematics provide.¹⁷ But the complex assumptions embedded in Pesendorfer’s mathematics are so poorly specified and, when interpretable, so bizarre that any operational challenge could be deflected by stating a divergence between the empirical challenge and the model’s assumptions. In a word, the model is non-operational. The exercise is barren. The paper is a great example of what does not count as science.

Unlike Pesendorfer’s model, in the real world: 1) populations are diverse (people are more than just “high” and “low” types); 2) people have freedom of association; they are not locked into a never-ending, randomized game that forces didactic associations based upon the distributions of garment designs in the population; 3) there are no clones, prisoners, nor cryogenic slaves available to satisfy a “matching technology”; 4) when it is important to them, people notice and remember the garments that others have worn; 5) “high” class women do not prefer to have other women (regardless of class) show up at social events identically attired; 6) consumers do not costlessly observe the contemporaneous purchase decisions of all other consumers; 7) consumers are constrained not only by the prices of fashion goods, but also by their incomes and the physical necessity of purchasing things other than fashion goods; 8) there is no single designer of fashion goods; 9) there are no temporal monopolies in designs; 10) advertising is not infinitely costly for old designs; 11) fashion designers do not have perfect foresight over future advertising expenditures nor their impact upon consumers; 12) advertising costs are not fixed in advance of their occurrence, and advertisers prefer to pay these costs later rather than sooner (*ceteris paribus*); 13) fashion designers are not limited to supplying only one design at a time; and finally 14) replicating fashion garments entails positive marginal costs.

It should be emphasized that our criticisms assume that

¹⁷ See Edward Chambers and Don Gordon (1966) for an exceptional example of how general equilibrium analysis can be used to gain real world insights; they investigated the impact of the wheat boom on the Canadian economy from 1901 to 1911.

Pesendorfer's model is for the purpose of providing operational propositions about specific phenomena. If it had alternative, albeit unstated, purposes our criticisms may be misplaced.¹⁸ Speculating on the "true" motivation behind an enterprise rather than its stated rationale is a sterile enterprise. We conclude that Pesendorfer's paper is inconsistent with observable reality. Milton Friedman (1953) famously maintained that in scientific inquiry it is legitimate to assert arbitrary assumptions about things that are not observable. Pesendorfer's error was to apply Friedman's dictum to things that are observable.

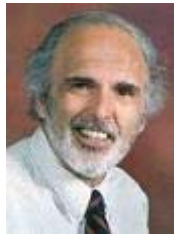
REFERENCES

- Baumol, William J.** 1990. Sir John Versus the Hicksians, or Theorist Malgré Lui?. *Journal of Economic Literature* 28(4): 1708-1715.
- Chambers, Edward J., and Donald F. Gordon** (1966). Primary Products and Economic Growth. *Journal of Political Economy* 74(4): 315-332.
- Davis, William L.** 2004. Preference Falsification in the Economics Profession. *Econ Journal Watch* 1(2): 359-368.
- Debreu, Gerard.** 1992. Random Walk and Life Philosophy. In *Eminent Economists: Their Life Philosophies*, ed. M Szenberg. New York: Cambridge University Press, 107-114.
- Friedman, Milton** 1953. The Methodology of Positive Economics. In his *Essays in Positive Economics*. Chicago: University of Chicago Press, 3-43.
- Gordon, Donald F.** 1955. Operational Propositions in Economics. *Journal of Political Economy* 63(2): 150-161.

¹⁸ Alternative justifications of economic modeling are characterized by skepticism and sarcasm: Axel Leijonhufvud (1973, 334) spoofs "There has been a great deal of debate in recent years over whether certain Econ modls and the associated belief-systems are best to be regarded as religious, folklore and mythology, philosophical and 'scientific,' or as sport and games." William Baumol (1990) suggests that the real purpose of some model building is model building, with economic phenomena serving as "hat racks" upon which to hang the models. Finally, Davis (2004) provides survey results indicating widespread skepticism about the official practices in the economics profession.

- Jones, E.L.** 1981. *The European Miracle*. Cambridge: Cambridge University Press.
- Leibenstein, Harvey.** 1968. Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. Reprinted in *Readings in Microeconomics*, ed. William Breit and Harold Hochman. Hinsdale, IL: Dryden Press, 123-139.
- Leijonhufvud, Axel.** 1973. Life Among the Econ. *Western Economic Journal* 11 (September): 327-337.
- Leontief, Wassily** 1971. Theoretical Assumptions and Nonobservable Facts. *American Economic Review* 61(1): 1-7.
- Pesendorfer, Wolfgang.** 1995. Design Innovations and Fashion Cycles. *American Economic Review* 85(4): 771-792.
- Samuelson, Paul A.** 1952. Economic Theory and Mathematics-An Appraisal. *American Economic Review. Papers and Proceedings of the Sixty-fourth Annual Meeting of The American Economic Association* 42(2): 56-66.
- Simmel, Georg.** 1957 [1904]. Fashion. *American Journal of Sociology*. 62(6): 541-558.

ABOUT THE AUTHORS



Philip R. P. Coelho is of professor of economics at Ball State University. He has published in *The American Economic Review*, *The Journal of Economic History*, *The Southern Economic Journal*, *Economic Inquiry* and other journals. Current research interests are the economic effects of morbid diseases, the biological basis of economic behavior, and ethics and behavior in business. His email address is: 00prcoelho@bsu.edu.



Daniel B. Klein is associate professor of economics at Santa Clara University, associate fellow at the Ratio Institute (Stockholm), and editor of *Econ Journal Watch*. He is the editor of [What Do Economists Contribute?](#) (NYU Press/Palgrave). His email address is: dklein@scu.edu.



James E. McClure (Ph.D., Purdue University, 1983) is Professor of Economics in the Miller College of Business at Ball State University (email address: jmcclure@bsu.edu). His publications appear in such journals such as: *Economic Inquiry*, *The Journal of Economic Behavior and Organization*, and *Southern Economic Journal*. His email address is: jmcclure@bsu.edu.

[GO TO RESPONSE BY WOLFGANG PESENDORFER](#)

Response to “Fashion Cycles in Economics”

WOLFGANG PESENDORFER*

THE PURPOSE OF MY PAPER “DESIGN INNOVATION AND FASHION Cycles” (*American Economic Review*, 1995) is to provide a model that captures two features of fashions.

1. Fashion goods are used as signaling/screening devices in social interactions.
2. Goods remain fashionable for a limited period of time and then go out of fashion only to be replaced by other fashionable goods.

The paper argues that fashion demand is cyclical *because* of the signaling role of fashion. Agents use fashion goods to signal their type—e.g., their wealth—and to screen the type of other agents. A fashion good is an effective signal as long as its price is high and only high types have an incentive to buy it. Over time producers will lower the price of the fashion good to sell the good to lower types. This leads to a degradation of the signaling value of the fashion. Eventually, there is room for a “new fashion”, i.e., another fashion item sold at a high price that separates high from low types.

In economic theory the role of models is to isolate the key aspects of the relevant reality. Stylized models serve this purpose by focusing attention on essential variables and by facilitating the analysis. To a large extent the criticism offered in the comment by Coelho, Klein, and McClure (CKM) reflects a misunderstanding of the role of economic models. Their criticisms often boil down to the assertion that the real world is more

* Department of Economics, Princeton University.

complex and that these complications should be reflected in the model. But adding realism is not a value *per se*. A more realistic, and hence more cumbersome, model that reproduces the same result obfuscates rather than clarifies the underlying insight.

THE PUZZLING FEATURES OF FASHION

Below I list two key aspects of consumer demand for fashion goods that my model seeks to explain.

Consumers pay a premium for fashionable labels, recognizable brands or fashionable designs. This premium cannot be explained by quality differences.

Prada sells Nylon bags that cost several hundred dollars. It is hard to imagine that consumers would pay the same amount for a bag of similar quality by a “no-name” producer. A key feature of Prada, and other producers of high-end fashion goods, is that their products are recognizable. People who buy a Prada bag can be sure that other fashion-conscious consumers recognize the brand of the bag and probably even the vintage of the design.

CKM seem to argue that there is no premium for fashionable brands and consumers do not pay for design *per se*. Hence, according to CKM consumers would pay the same amount for a Prada bag and a similar quality bag by an unknown producer.

Desirable designs go out of fashion only to be replaced by new desirable designs. Consumer demand for fashion is surprisingly correlated.

Every observer of fashion notes that fashions “change”, i.e., the demand for a fashionable product will be high for a limited period of time and then will drop. Periodically new fashions emerge and consumer demand tends to be highest for the latest designs.

A naïve theory that tries to explain this phenomenon with a taste for variety falls short because it does not explain the correlation across consumers. When a new fashion arrives all consumers seem to switch their tastes simultaneously. This simultaneity cannot be explained by an individual taste for variety. Rather it suggests that the demand for fashion is a “social” phenomenon, i.e., the demand for fashion has something to do with social interactions.

A BRIEF SUMMARY OF THE THEORY

People often interact with individuals who they do not know very well. Not all interactions are equally profitable and individuals must decide whether or not to “invest” in an interaction without knowing the exact type of the potential partner. A type may refer to the income, the education status or the intelligence of a person. In this situation it may be beneficial if a potential partner can signal his or her attractiveness. The role of fashion in this context is to facilitate this signaling and help sort individuals according to types.

The paper offers a model of this role of fashion. Since we want to focus on the social role of fashion, it makes sense to assume that the fashion good has no intrinsic value to the consumer. The point is that *even if* the fashion good has no intrinsic value it may be useful as a signaling device in social interactions.

The paper assumes there are two types, “high” and “low”. Agents derive a benefit from interacting with one other agent. This benefit depends on the types of the two individuals. Types are complementary so that positive assortative matching is socially optimal.¹ Since the type of an individual is unobservable, the socially optimal matching cannot be implemented. In the benchmark case without fashion, each agent is simply matched with a randomly drawn individual. Fashion is modeled as an (indivisible) item that is visible but otherwise useless. Suppose that some agents buy the fashion item and others do not. It is now possible to sort agents into those who use the fashion item and those who do not. Because high-type individuals are willing to pay more than low type individuals to

¹ This means that it is socially optimal to match high types with high types and low types with low types.

improve their chance of meeting other high types the fashion facilitates more efficient matches.

The model assumes that agents who use the fashion item are matched with other agents who use the fashion item and non-users are matched with other non-users. This matching process is incentive compatible because, in equilibrium, agents who own the fashion item are more likely to be high types. The design-users prefer to be matched with other design-users and will reject being matched with non-users. Hence, the assumed matching technology can be thought of as the stable outcome of a decentralized, voluntary matching process.

A key ingredient of the theory is the unobservability of types. If types are observable there can be no social value to fashion as described in the model. Hence, the theory suggests that we should see more use of fashion in urban environments where anonymous interactions are more common than in rural areas where anonymous interactions are rare. In rural areas individuals may have pretty good information about the types of potential matches and therefore have less demand for signaling devices to aid the matching process.

To illustrate the dynamics of the model suppose there is an established fashion that separates high and low types. Suppose high types have purchased the fashion item and low types do not find it worthwhile to buy the fashion item at the current (high) price. Consider the incentives of a firm selling the fashion good. If the marginal cost of producing the fashion item is small then the firm has an incentive to lower the price and sell the item to low types. Low types will use the fashion to improve their chance of meeting high types. But, when all consumers use the fashion, it can no longer separate types. As a result, there is demand for a new fashion. A new design, if introduced at a sufficiently high price, will again separate types.

The theory argues that goods go out of fashion because they cease to be effective at separating types. The model assumes a monopolist seller of the design. Because fashion items are durable the monopolist will lower the price of the item over time. Once the item is owned by a sufficient number of consumers there is room for a new fashion. At that point the monopolist will create a new design and a new cycle begins.

One key assumption is that the fashion good is relatively cheap to produce (i.e., a low marginal cost). This is a reasonable assumption for Prada handbags made of nylon but a less plausible assumption for jewelry or high-end watches. In the latter case, the good may still be used to as a signaling/screening device. However, if the marginal cost of producing the good is high the seller may not lower the price sufficiently for low types to

buy the good. As a result the signaling/screening role of fashion may not lead to fashion cycles for goods with a high marginal cost. Put differently, the theory suggests that we should see fewer fashion changes for products with a high marginal cost such as high-end watches than for products with low marginal costs such as nylon handbags.

CKM observe that in reality there is more than one producer of fashion, and hence the assumption of a monopoly is unrealistic. However, the point is that *even if* the designer has monopoly power over the design (and imitation is impossible) we can expect degradation of the signaling value of the design over time. Hence, even when there is a monopoly we can expect design changes and fashion cycles. If designs can be imitated then we would expect faster fashion cycles but the driving force behind the cycle would be the same: as designs spread they become less valuable.

In reality, designers do have some monopoly power in the sense that they control the label. It is illegal for other designers to *exactly* copy a Prada bag or to use the Prada label. In that sense, Prada has some property rights over its designs.

ADDRESSING SPECIFIC CRITICISMS

In this section, I address some of the specific concerns raised by CKM.

No Intrinsic Value of Fashion Goods

CKM write: “*Unlike real-world garments Pesendorfer’s ‘design’ does not comfort, protect, warm, nor beautify; neither does it generate prestige or ostentation*” (CKM, 438).

The point of my paper is to analyze aspects of fashion that differ from standard goods. Of course, clothing keeps people warm just like orange juice quenches thirst. But what is interesting about clothing is that it sometimes has a social role in addition to its more traditional role. The paper focuses on this social role and therefore it makes sense to abstract from other functions of the fashionable item.

The paper shows that *even if* there is no intrinsic value to a fashion good it may have value for consumers because it helps them in social interactions. The model tries to isolate the difference between fashion and orange juice. An unrealistic but useful simplifying assumption is to assume that fashion has no intrinsic value.

It is clear that certain ways of introducing an intrinsic value to fashion goods will not alter the results. For example, a type-independent intrinsic utility would not qualitatively change the results. However, it is not clear that *every* way of introducing intrinsic value would leave the analysis unchanged. The purpose of adding intrinsic value to the model would be to figure out when it alters the analysis and when it does not. CKM do not explain what, if any, effect they expect from the introduction of intrinsic value.

I disagree with the assertion that fashion in my model does not generate prestige or ostentation. In my model, people own fashionable items in order to affect the pool of agents they interact with. Those who own the fashion item are more likely to be high types and therefore are more likely to meet other high types. Far from precluding prestige or ostentation, the model provides a specific theory of what one could mean by prestige (probability of being a high type) and explains why people care about prestige (increased probability of being matched with other high types).

A related complaint in CKM is that the model is not specific in what is meant by fashion. CKM write: *"Pesendorfer uses the terms 'fashion', 'garment' and 'design' interchangeably because, in the model, the all mean simply the ticket to mix with other ticket holders"* (CKM, 438). Again, the authors express a distaste for abstraction and a desire for nuanced realism. For my results it is irrelevant whether the fashionable item is a garment or a designer handbag. Providing a unified framework for analyzing fashion cycles without having to address the specifics of particular industries is one of the main contributions of my paper.

Matching

CKM object to the matching process as a model of social interactions. They refer to it as a *"compulsory, never ending matching process"* (438) and suggest that it describes a world of *"forced association"* (442).

The matching technology sorts people by the fashion they use. If there is one fashion design then each user of the fashion item is matched with another user while non-users are matched with other non-users.

This matching technology is simply an abstract way to capture decentralized interactions. As is typical in economic theory, the voluntary aspect of the interactions is captured by the equilibrium. *In equilibrium*, people who use the fashion prefer to be matched with other people who use the fashion and would object to being matched with people who do not use the fashion. The reason is that users are more likely to be high types than non-users. Therefore, in equilibrium, the “forced” matches are all voluntary. In fact, the “forced” matches are the only voluntary matches possible in equilibrium.

CKM could level the same criticism at the standard Walrasian model of competition. The competitive model “forces” agents to trade at a given price. A literal interpretation of the model might argue that the model describes a planned economy and not a market economy because agents are not free to choose the terms of trade. This critique misses the point that (in a competitive, thick markets environment) *equilibrium* prices are the best terms an agent could hope for. If a consumer proposed a price that is more favorable to him than the equilibrium price he would be unable to find a trading partner. Hence, *in equilibrium*, the competitive model captures the outcome of voluntary and decentralized exchange.

CKM also object to the following assumption on the matching process. Suppose all but one agents use a design. It is assumed that the single non-user is matched with a low type. To motivate the assumption it is useful to consider what happens when there is a small fraction of non-users. In that case, *in equilibrium* the non-users are sure to be low types. As a result, an agent who does not use the design is sure to be matched with a low type. Hence, the assumption that a single non-user is matched with a low type amounts to assuming a continuity of equilibrium payoffs at the point where the design reaches full market penetration. The assumption can be justified if there are some low types who are committed to never using the design.

An alternative would be to assume that a single non-user is not matched and therefore receives a lower payoff than if he were matched with a low type. This would lead to a discontinuous jump up in the reservation price for the fashion item at the point where the item has full market penetration. I conjecture that even with this discontinuity the results of the paper would be qualitatively unchanged.

CKM seem to suggest (441) that in the case where all but one agents use the design, the single non-user should be matched with a random *user* of

the design. However, users would *object to* being matched with a non-user since they will (reasonably) believe the non-user is less likely to be a high type than a user. Therefore, it would be unreasonable to match a single non-user with a random user.

Demand Function

The analysis of the demand for fashion in a static setting asks how much agents would be willing to pay if the output was q units. The function f describes the maximum that agents would be willing to pay for q units. Hence, f can be interpreted as the market's willingness to pay or the inverse demand function.

CKM are correct when they argue that f is not the same as a standard demand function. There can be no immediate analogue of a demand function because the value of the fashion good is entirely derived from its role facilitating better matches.

CKM point out that if we simply set a price then there may be multiple demands consistent with this price. This is correct but irrelevant. The producer can pick the price *and* the quantity he chooses to supply. The function f describes the possible price/quantity choices that are feasible for a (monopoly) producer in the static setting.

The Importance of Design

CKM (449) assert that “...no quasi-monopoly in fashion design exists. What principally allows garment producers to price their products at a premium to ‘ordinary’ garments is their reputation for producing superior garments, superior in a number of production characteristics.” Further CKM claim (450): “People in the fashion trade realize that the marginal cost of a particular design is trivial. Consequently, consumers do not pay for design per se.”

These assertions are stated without evidence and do not seem plausible. For example, when a consumer buys a \$665 Prada handbag made of Nylon,² what is she paying for if not the design? If CKM's claim were

² Available at Neiman Marcus online. See:

<http://www.neimanmarcus.com/store/catalog/prod.jhtml?cmCat=search&itemId=prod15690701>

true then a bag by a no-name producer similar in quality to the above mentioned Prada bag should sell for a similar price. I find this very implausible.

High-end fashion producers are known for their designers. Many of these companies (D&G, Armani, Prada, etc) carry the name of their star designer and most of the marketing effort is focused on selling a “look”. For CKM this must be a very surprising coincidence. Since in their view fashion is all about the quality of the garment the celebrities of the fashion world should be the high-quality manufacturers and not the designers. If CKM’s view of the fashion world were correct we should expect the quality of the manufacturing to be the central focus of the fashion market. As is the case for other consumer durables, we would expect the marketing to focus on functionality and durability rather than looks.

Notice that this focus on selling a “look” is exactly what one would expect if fashion is used as a signaling device. After all, the appearance of a garment is the one feature that is observable both to the consumer *and to the observer*.

Fashion houses own a “label” or a brand name that is valuable. My model suggests an explanation why a label can trade at a premium even if the no-label substitutes are of similar quality. The reason is that consumers use the label to signal something about their type. Note that entry in this market is difficult because any newcomer must solve a coordination problem. To establish a new “label” it is not sufficient to convince a single consumer to switch. Rather the fashion house must convince a whole population to adopt the new label as an accepted signaling/screening device. Moreover, offering the good at a low price may not be a good strategy for an entrant because at a low price the fashion is ineffective as a signaling/screening device.

CONCLUSION

In the paper, “Design Innovations and Fashion Cycles”, I argue that fashion goods have a signaling/screening role that helps explain fashion cycles. The point of the model is to develop a framework that illustrates the mechanism by which the signaling/screening role of fashion leads to fashion cycles.

In their comment, CKM argue that many of the simplifying assumptions in my model are unrealistic. However, they offer no arguments why more realistic (and more cumbersome) counterparts would overturn the basic result of the paper. In this response, I have argued that many of the simplifying assumptions are well justified and that the mechanism by which the signaling/screening role of fashion leads to fashion cycles is fairly robust.

REFERENCES

- Coelho, Philip R. P., Daniel B. Klein, and James E. McClure.** 2004. Fashion Cycles in Economics. *Econ Journal Watch* 1(3): 437-454.
- Pesendorfer, Wolfgang.** 1995. Design Innovations and Fashion Cycles. *American Economic Review* 85(4): 771-792.

ABOUT THE AUTHOR

Wolfgang Pesendorfer is Professor of Economics at Princeton University. His research interests include game theory, political economy, and decision theory. His e-mail address is pesendor@princeton.edu.

[GO TO COMMENT BY COELHO ET AL.](#)

Rate of Economic Growth, Level of Development, and Income Inequality: Rejoinder to the Reply by Edwards and McGuirk

JIH Y. CHANG* AND RATI RAM**

**CONTINUATION OF EXCHANGE BETWEEN JEFF EDWARDS AND
ANYA MCGUIRK, AND JIH Y. CHANG AND RATI RAM FROM THE
AUGUST 2004 ISSUE OF *EJW*.**

[Edwards and McGuirk's Comment on Chang and Ram](#)
[Chang and Ram's Response](#)
[Edwards and McGuirk's Reply](#)

WE APPRECIATE THE OPENING PARAGRAPH OF THE REPLY BY Edwards and McGuirk (2004b). However, the Reply contains largely irrelevant or inconsequential statements. To place this exchange in perspective, we recall that our original paper (Chang and Ram 2000) suggested, on the basis of estimates of a Kuznets-quadratic from cross-country data, that high-growth economies are likely to experience lower income inequality at any given income level. In addition to some secondary matters, Edwards and McGuirk (2004a) claimed that our conclusion does not hold if two regional intercept dummies are included in the regression model. We pointed out (Chang and Ram 2004) that Edwards-McGuirk's

* Financial Aid Office, Illinois State University.

** Department of Economics, Illinois State University.

claim was simply inaccurate because, contrary to their own basic point, the predicted inequality numbers were derived by ignoring parameters for the regional dummies. We also showed that even when the two regional dummies are included, one observes essentially the same pattern as indicated by Chang and Ram (2000) if accurate predicted values of inequality are used. The long Reply (Edwards and McGuirk 2004b) provides no meaningful basis for altering the aforesaid position. The contents of the Reply appear largely irrelevant to the main point or are inconsequential or inaccurate. It might have been reasonable for them to “step back and give some perspective on” their comment if they focused on the core of the exchange and dealt with secondary matters briefly in a transparent manner. However, most of the Reply deals with secondary aspects in a diffuse style, which makes it difficult for us to formulate a structured response. The following paragraphs offer brief observations on a few aspects that we are able to comprehend in the Reply.

1. Their observation about being “skeptical of, and disappointed with, the current state of published empirical work in economics” (244) has little relevance to the substance of the exchange. Irrespective of the basis of their comment, they should deal with the main point of our reply (Chang and Ram 2004). If they carry any general dissatisfaction or disappointment about the state of empirical research in economics, its expression is irrelevant to this exchange. To the extent their reservations can be cogently articulated, these belong in a more general paper about the “state of published empirical work in economics.” Their Reply should deal primarily with the point whether Chang and Ram (2004) are right in saying that the predicted inequality numbers used by Edwards and McGuirk (2004a) were wrong for many countries and that accurate predicted values from even their own model yield largely the same pattern as noted by Chang and Ram (2000). It is inappropriate to introduce more general issues about the “state of empirical research” or to bring in views of scholars like Leontief, Pagan and Spanos on those general aspects.

2. Their observations on page 246 about what they were doing in the comment seem like an afterthought. At any rate, most of this part is redundant in response to the main point of our reply. While one wishes that these assumptions were “easily testable,” listing of the standard assumptions of the linear regression model seems particularly pointless since these can be found in almost any elementary econometrics text. We also believe that the word “forged” from Pagan's quote, stated (246) as a possibility relative to our results, might be more applicable to a procedure

like that followed by Edwards and McGuirk in which the specification (and the set of estimates) is “tortured until it confesses.”

3. The further discussion (Edwards and McGuirk 2004b, 247-248) about the choice of regional dummies is also irrelevant. Although the dangers of dropping “insignificant” variables are well known, there is no need to make an effort to justify their choices when we have shown (Chang and Ram 2004) that even with the inclusion of their two regional dummies, Chang-Ram (2000) conclusion is not affected if accurate values of predicted inequality are used.

4. Their statement (249) about what they did in the comment (Edwards and McGuirk 2004a) is wrong. They generated predictions inaccurately by ignoring parameter estimates for the regional dummies and (amazingly) assuming “the countries are all in the control group” (Edwards and McGuirk 2004a, 230). While it is obvious that their procedure would generate accurate predictions for the base group, they can derive no comfort from that since the predicted values for all “nonbase” countries were wrong. It is surprising that Edwards and McGuirk are hesitant to acknowledge even this obvious and major error, and are trying to obscure it. As a minor aspect, what we said on this point (Chang and Ram 2004, p. 239) was that the comment ignored parameters for regional dummies, and not that they did not model regional differences.

5. The “another attempt to illustrate the differences in predictions” (Edwards and McGuirk 2004b, 250) is misleading. It is not correct to compare high- and low-growth inequalities for the “base” countries separately from those for “CA” and “E” regions. Since the estimates are derived from the combined sample, predicted inequalities for high- and low-growth cases *must* be compared for the entire group of 48. The blue and red “smooth” curves in their Figure 1 (250) are a distortion. The related statements about there being “little difference in the inequality-GDP relationship between . . . high- and . . . low-growth countries,” the differences being “economically insignificant,” and their model not predicting “that higher growth countries have lower inequality at all income levels,” are also misleading. These refer only to the base group and are inaccurate. The only valid and meaningful comparison is between predicted values of inequality for *all* high-growth and *all* low-growth cases in the sample. These predicted values are shown in Table 2 of our reply (Chang and Ram, 2004, 240) and plotted in our Figure 1 (241) which supports the conclusion stated in our original paper (Chang and Ram 2000). If accurate, every point in their Figure 1 (Edwards and McGuirk, 2004b, 250) must be

identical with that of our aforesaid Figure 1; the separation into regions is misleading and seems intended to obscure the failure of their comment.

6. The statements below their Figure 2 (251) are inconsequential or inaccurate. It is obvious from their own exposition that the two regional dummies capture to a substantial extent the high- and low-growth dichotomy in the sample. At any rate, we have shown that the correct predictions even from models that include the two regional dummies used by Edwards and McGuirk yield the conclusion that high-growth countries have lower inequality.

7. The long discussion about Kuznets hypothesis (Edwards and McGuirk 2004b, 252-255) is a distraction and misleads the reader from the main point. Edwards and McGuirk (2004b, 252) acknowledge that this is a “minor focus” for us, but still devote more space to this than to the core points. The futility of this part of their Reply is indicated by several other considerations also. First, if they look at their own Figure 1 (Edwards and McGuirk 2004b, 250), they will find an unmistakable Kuznets-U in each of the four smooth curves, which, of course, are not relevant to a comparison of the inequalities in high- and low-growth cases. Second, the scatter in their Figure 3 (252) is incomplete since it includes only 48 countries, and cannot be used to judge the presence of a Kuznets pattern in the sample. Even when the hypothesis holds in the entire sample, one can always find a subset where it does not hold. We have shown (Chang and Ram, 2004, 237) that even when two regional dummies are included, there is strong evidence of a Kuznets-U in the full sample. Although we postulate parametric variations in high- and low-growth cases, the specification used by Edwards and McGuirk (2004b, 253) is less appropriate than ours since Kuznets-hypothesis should be explored on the basis of what might be called “averaged” parameters for the entire sample. It is pointless to try to explore the hypothesis in subsamples. It seems Edwards and McGuirk (2004b) do not realize that our consideration of low-growth and high-growth samples was meant *only* to show that the evidence on the hypothesis depends on the data and not the quadratic model. It is difficult to see what could possibly be gained by redoing our exercise after including regional dummies. Incidentally, contrary to what Edwards and McGuirk (2004b, 255) state, their different findings on Kuznets hypothesis for low- and high-growth cases are obviously detrimental to the claim that there are no significant parametric differences between the two groups.

8. The part on “final aspect of (our) complaints” is also inaccurate or inconsequential. We have addressed the “first aspect” in the foregoing paragraphs. To repeat, even using the model that Edwards and McGuirk

find “statistically valid,” accurate numbers for predicted inequality refute the claim made by them in the comment and support the Chang-Ram (2000) conclusion. We have also noted that the “fatal flaw” in their comment was to use inaccurate values of predicted inequality for many countries. The “second aspect” is based on a false premise. There is *not* an “overall lack of significance of most of the variables in the model.” On the contrary, our EDCC estimates (Chang and Ram, 2000, 792), on the basis of which Edwards and McGuirk proceed further, show that, despite the limited sample size and considerable collinearity, each income term is significant at least at the 10% level, and the three high-growth dummies are jointly significant at the 6% level. Even in the specification that includes regional dummies, the three high-growth dummies are jointly significant at better than the 10% level. These are reasonable significance levels, and the premise on which Edwards and McGuirk proceed is false.

In addition to being based on a false premise, introduction of confidence intervals by Edwards and McGuirk is not meaningful for several other reasons also. First, if confidence intervals were really an important consideration relative to our paper, their comment (2004a) should have focused on these instead of complaining about the “Kuznets curveball missing the regional strike zone.” Like the bulk of the Reply, confidence intervals are an afterthought, and amount to improperly writing another comment. Second, standard practice in the profession is to base inferences about the significance of the estimates and the parametric differences on the point estimates and the related standard errors. We have already noted that the three high-growth dummies in our (2000) estimates, which Edwards and McGuirk use for confidence intervals, are jointly significant at the fairly stringent 6% level and each income term is significant at least at the 10% level. Structure of confidence intervals would just reflect the point estimates and the standard errors and can shed *no* light on the “significance” of the differences in predicted inequalities for the two groups. Third, by choosing to work with 95% confidence intervals, Edwards and McGuirk are apparently looking for significance at the 5% level. It is not evident where the sanctity of the 5% level comes from. It is just one of the conventional levels, and the 10% level may be more appropriate in this case. Fourth, if Edwards and McGuirk consider 5% as the only acceptable level for judging “significance” or “confidence,” it is evident that our estimates (Chang and Ram 2000, 792) do not meet that criterion. Instead of writing 29 pages of the comment and the Reply, Edwards and McGuirk could just have written a page suggesting rejection of our conclusions due to lack of significance at the 5% level. Fifth, aside from everything else, and

even assuming that the confidence intervals are accurate, these support our (2000, 2004) position that high-growth economies are likely to experience lower inequality. It can be seen from their Figure 4 (257) that, like the predictions based on the point estimates, each bound of the 95% confidence interval for high-growth cases lies below the corresponding bound for the low-growth cases at almost every income level. Thus the pattern indicated by the predicted values based on the point estimates is reinforced by confidence intervals. Judging, without any criterion, “width” of the confidence intervals, or considering the location of predicted values relative to the intervals, can generate no inference about the “degree of confidence” or significance of the difference between the inequality profiles in the two groups. By introducing confidence intervals and mentioning the aforesaid characteristics of the intervals, Edwards and McGuirk seem to be only trying to obscure the failure of their comment. As an aside, it is interesting to note that Edwards and McGuirk work with confidence intervals based on our estimates and not on their “statistically valid” model. Perhaps the confidence intervals based on their model did not look as “good” despite being “almost as large as those of Chang and Ram” (257).

We also make a few remarks about their “thinking” articulated near the bottom of page 255. As we tried to explain in our reply (Chang and Ram 2004), it is *not* true that if the variables are unrelated, estimating a quadratic specification would indicate “some sort of quadratic relationship.” The reasoning is spurious. Following that argument, one can also say that if the variables are unrelated, estimating a linear model would indicate “some sort of a linear relationship” and estimating a cubic (or a logarithmic function) would indicate “some sort of a cubic (or logarithmic) relationship.” It should also be noted that, contrary to what Edwards and McGuirk seem to suggest, patterns of predicted inequality are not dependent on a quadratic specification. These should show in any reasonable functional form. We used the Kuznets-quadratic because it is a widely-adopted model for studying the relation between income and inequality.

10. While Edwards and McGuirk claim to worry much about “statistical adequacy” of the models and “reliability of inference,” they seem to overlook the point that consideration of an appropriate level of significance is an important part of statistical inference. Given the extensive discussion in the literature about the nexus between income and inequality, it is probably not true that the hypothesis of income (level of development) having no association with inequality is so strong as to require overwhelming evidence for its rejection. On the contrary, a reasonable

approach would probably be to reject the null on the basis of “considerable,” and not overwhelming, evidence, and a significance level of 25% (or even less stringent) might be more appropriate than the conventional levels. Such a thought renders even less meaningful their “second aspect,” the remarks about “overall lack of significance of most variables,” and introduction of 95% confidence intervals.

We conclude by saying that, while we appreciate the tone of the opening paragraph of Edwards and McGuirk (2004b), nothing in their Reply mitigates the failure of the main argument of their comment (Edwards and McGuirk 2004a) that inclusion of two regional dummies alters the Chang-Ram (2000) conclusion. We showed in our Reply (Chang and Ram 2004) that (a) the predicted inequality numbers used by Edwards and McGuirk (2004a) in the comment were wrong in a major way, and (b) accurate predicted values even from their “statistically adequate” model yield the same pattern as stated in our original paper (Chang and Ram 2000). On the first point, Edwards and McGuirk seem to admit the error, although only indirectly and hesitantly. On the second point, the only relevant parts in the long Reply are expressed through Figure 1 (250) and Figure 4 (257). We have shown that the “smooth curves” in Figure 1 are a distortion and are irrelevant to the issue. Figure 4 is based on a false premise, is clearly an afterthought, chooses an arbitrary confidence level, is not consistent with standard econometric practice for judging “significance” of differences in the parameters (and the related fitted values), and can yield no information about the significance of the differences in the inequality profiles for the two groups. Moreover, even if one were to make some use of 95% confidence intervals for shedding light on the main point, an appropriate interpretation of the confidence intervals reinforces our conclusion that high-growth economies are likely to experience lower income inequality.

Despite the failure of the main argument of Edwards and McGuirk, we do not claim that our conclusion is infallible. We are conscious of the hazards of drawing strong conclusions from studies of this type, and made it clear (Chang and Ram, 2000, 795) that the stated inference is subject “to the caveats appropriate for studies such as this one, which work with simple models and cross-sectional data from samples of a modest size.” The difficulties in such contexts, particularly the one about drawing policy-relevant inferences from cross-country data, are well known. However, the points made in the comment and the Reply by Edwards and McGuirk (2004a, 2004b) have almost no bearing on those difficulties.

REFERENCES

- Chang, Jih Y., and Rati Ram.** 2000. Level of Development, Rate of Economic Growth, and Income Inequality. *Economic Development and Cultural Change* 48(4):787-799.
- Chang, Jih Y., and Rati Ram.** 2004. Response to Edwards and McGuirk: Income Level, Economic Growth, and Inequality: Flawed Methodology and Inaccurate Inference. *Econ Journal Watch* 1(2):235-243.
- Edwards, Jeff, and Anya McGuirk.** 2004a. Kuznets Curveball: Missing the Regional Strike Zone. *Econ Journal Watch* 1(2): 222-234.
- Edwards, Jeff, and Anya McGuirk.** 2004b. Reply to Chang and Ram: Statistical Adequacy and the Reliability of Inference. *Econ Journal Watch* 1(2): 244-259.

ABOUT THE AUTHORS

JIH Y. CHANG is Assistant Director in the Financial Aid Office, at Illinois State University, Normal, IL.

RATI RAM is Distinguished Professor of Economics at Illinois State University, Normal, IL.

Character Issues

Scholasticism versus Pietism: The Battle for the Soul of Economics

ROBERT H. NELSON*

[Abstract, Keywords, JEL Codes](#)

[E]conomic ideas—ranging from new but unelaborated concepts through isolated propositions about causality, all the way to full-blown theories—arise in the highbrow part of the economics profession and then diffuse first within the profession and then sometimes outside it to journalists, bureaucrats, politicians, and other citizens.

—Robert Solow (1989, 75)

SINCE THE 1930S ECONOMICS HAS BEEN FREQUENTLY CRITICIZED for being irrelevant and, when possibly relevant, unintelligible. To characterize this failing, critics have often used the term “scholasticism.” Indeed, that term is generally used, as *Webster’s Revised Unabridged* (1998) holds, to mean “characterized by excessive subtilty, or needlessly minute subdivisions; pedantic; formal.” Such scholasticism, moreover, is generally a product of a certain *structure* of discourse and exploration: a top-down hierarchy based on a public means of support. Here I distinguish between *topical scholasticism*, meaning irrelevancy and pedantry, and *structural scholasticism*, meaning a social structure based on hierarchical validation and involuntary financing. Topical scholasticism usually depends on structural scholasticism, but structural scholasticism can in principle strive to avoid being scholastic; it can strive to be oriented toward meaningful issues and relevant public

* School of Public Policy, University of Maryland.

discourse. That goal would likely be better achieved, however, by undoing the scholastic structure.

KNOWING GOD: SCHOLASTICISM VERSUS PIETISM

How does a person know God and his ways? In the history of western theology there have been two opposing views. In one view—which I call the “scholastic” view—a church hierarchy interprets the ways of God to the faithful. An official priesthood transmits the results of a long history of internal church discussions and debates. In this process, great legitimacy is given to having an exhaustive knowledge of past religious documents, historic writings, and other relevant materials. The most skilled in the use of these materials often claim that the results demonstrate human “rationality” at its highest levels. In the Middle Ages, the Roman Catholic Church conducted its internal discussions in Latin, thus precluding the possibility of participation or comprehension by the ordinary people. Paul Tillich, perhaps the most distinguished American theologian of the twentieth century, has said of this tradition that “the Roman system is a system of divine-human management, represented and actualised by ecclesiastical management” (Tillich 1967, 228).

In a contrasting view—which I call here the “pietistic” view—there is a more direct relationship between the individual and God. Protestantism in general preaches that salvation is “by faith alone”—without any essential intermediary role for a church hierarchy. In the sixteenth century the Protestant Reformation disbanded the large church land holdings and other properties and abolished the priesthood of the Catholic Church. Protestantism instead preached a “priesthood of all believers”—every person should be equally devout and committed to the dedicated pursuit of God’s truths.

Many of the knowledge claims of the Catholic priesthood were dismissed by the Protestant Reformers as the self-serving manipulations—empty “scholastic” exercises—of a corrupt church that sought thereby to maintain a religious monopoly. Protestants were instead encouraged to study on their own the original source materials, especially the bible, as a central element of their religious life. The absence of an intermediate church hierarchy led to a new individual intensity in the relationship of each Protestant faithful with God. As Tillich wrote, it is a “person-to-person

relationship” which is “unconditional” and in which “one is not a bit nearer to God if one does more for the church” (Tillich 1967, 229).

The division of the [church] faculty into “experts” is a very unwholesome state of affairs, where the New Testament man tells me that I cannot discuss a certain problem because I am not an expert, or I say that I cannot discuss a matter because I am not an expert in Old or New Testament. Insofar as we all do this, we are sinning against the original meaning of Luther’s attempt [to reform the Christian methods of doing theology]. ... These are very real problems [even] today, and students can do a great deal about them by refusing to let their professors be merely “experts.” (Tillich 1967, 244-245)

Yet, even the Protestant Reformers found it necessary to develop a systematic body of theological writings. Both Luther and John Calvin wrote many volumes of biblical and other theological exegesis and both they and their followers sometimes were dogmatic in insisting that the faithful subscribe to these views. Although it was in many ways inconsistent with the main reform thrust of Protestantism, over time such writings could easily become an official orthodoxy, and scholastic tendencies arose within Protestantism as well. Some branches of Protestantism such as the Anglicans in England maintained a hierarchical structure of church practice and teaching. At the same time, there were some important reform movements within the Roman Catholic Church that attacked the sterility of existing scholastic methods and tendencies. Given the absence of a single source of church authority, however, it was easier for powerful new reform movements to arise among the Protestant faithful, which happened on a frequent basis in the next few centuries.

By the middle of the seventeenth century many of the established churches of Protestantism were showing signs of what the historian of theology Gerald Cragg calls a new “aridity of theology” and a “lifeless and unbending orthodoxy.” The radical Puritans in England offered one challenge. In Germany a movement of religious revival, Pietism, was led in the later part of the century by Phillipp Jakob Spener. Critical of developments in the Lutheran church, Pietism elevated the role of the laity in the life of the church and emphasized that religion should not be a matter of the learning of formal theology, but must be lived on a daily basis—that “Christianity was not an intricate system of abstruse doctrines

but the practice of a transforming way of life” (Cragg, 1970, 101). The pietist faithful were committed to their own intensive study of the Scriptures in order that they could commit their lives better to the teachings of Jesus. In this manner, pietism “broke the paralyzing hold of Lutheran scholasticism” (Cragg 1970, 103). As Cragg comments, “in its resistance to control of religious opinion, Pietism represented an outspoken assertion of individual rights in the face of the entrenched prerogatives of the civil rulers” and their religious backers (Cragg 1970, 105). Pietism stands in contrast to scholasticism in relying on the common sense judgments and actions of the individual person, as against the collective authority of any established clerical hierarchy.

RELIGIOUS ENLIGHTENMENT AND ECONOMIC ENLIGHTENMENT

Today, there is an analogous tension. Instead of knowledge of God, the issue is economic enlightenment. The members of the economics profession who practice and affirm the formalistic genres of economics are like the scholastics of old. The economics profession is hierarchical. It works by internal processes based on a well established ranking of prestige in the application of “rational” methods. The most authoritative economists have formed their own exclusive society, validate each other’s station, replicate their kind in PhD programs, and maintain control over the whole field by means of appointments, publications, and so forth. Professional economists communicate in a language of mathematics—the “Latin” of our time—that similarly excludes ordinary people. A true understanding of economic processes is said by economic professionals to be possible only within a framework of formal economic analysis.

The official keepers of the faith of the new “church” of this modern scholasticism are found in the leading university departments of economics and at the “top” journals (which are almost always edited at one of the elite departments). Other supporting institutions of the modern scholastic church of economics include the American Economic Association (whose leadership is usually from the prestige departments); the Economics Division of the National Science Foundation (which relies heavily on reviewers from these same departments); the National Bureau of Economic Research; the National Academy of Sciences; and still others who act to

filter out any “heretical” tendencies. The economic miscreants are no longer burned at the stake, but they are still effectively excluded from power and influence within the economic priesthood.

The economic priesthood is mostly a product of the twentieth century and the rise of the professional classes. Adam Smith—of Protestant Scotland—worked outside the academic world and was writing in the eighteenth century for the full literate population of Britain. Indeed, Smith saw the universities of his time as an obstacle to the advancement of intellectual understanding.

The improvements which, in modern times, have been made in several different branches of philosophy have not, the greater part of them, been made in universities, though some no doubt have. The greater part of universities have not even been very forward to adopt those improvements after they were made; and several of those learned societies have chosen to remain, for a long time, the sanctuaries in which exploded systems and obsolete prejudices found shelter and protection after they had been hunted out of every other corner of the world. (Smith [1776], 1937)

In reading *The Wealth of Nations*, ordinary people could well enough understand that a free market system would affirm their liberties and ensure their future prosperity. No official body of priests was required to validate this message. Today, there are heirs to Adam Smith, although seldom found in the highest ranks of the economics profession. A new body of economic writings is found in “pietistic” organizations that depart from the official “Latin” of professional economics.

Organizations such as the Foundation for Economic Education, the Cato Institute, and the Institute of Economic Affairs do not appeal to priestly authority, and do not lobby the powers that be to impose their doctrines on others (they oppose the government production of schooling). They appeal to the common understandings of interested lay observers.¹

¹ The three examples given are libertarian examples, and certainly other examples from other ideological quarters could be given. However, I would argue that the more mainstream “liberal” and conservative organizations and periodicals tend to approach issues politically rather than economically; they appeal more to the sporting nature of the political contest. As for the leftist periodicals outside the center, I would argue that their emphasis on elucidating economic principles is very weak. Thus, in endeavors of economic pietism, I see a certain prominence to libertarian ideas.

They would dispute any suggestion of an exclusive monopoly of economic knowledge possessed by a limited body of professional “experts.” They call on ordinary people to ponder and assess the merits of rival claims to economic enlightenment.

These issues concern the character of economics, raising such questions as the following: How technical is economic knowledge? Is valid economic knowledge mainly developed through the workings of a priestly or professional hierarchy? When economists appeal to a scientific status, is this mainly a means of asserting a special claim to a priestly control over the development and use of economic knowledge? How much real addition to economic knowledge has been achieved as a result of the professionalization of economic inquiry in the twentieth century? Or perhaps has this professionalization instead inhibited the advance of greater economic understanding (as was the case in some earlier eras of scholastic inquiry)? Do we in fact know much more than Adam Smith about the economy, or has the greatest change been in the manner of presentation of much the same economic knowledge? Might the scholastic discourse even tend to eclipse or avoid truths that Adam Smith long ago already appreciated? Are the scholastic preoccupations harmful movements for the overall advance of economic knowledge?

To address these questions systematically would require an effort that would extend well beyond the scope of any one article—or perhaps any one book.² I propose instead to examine these tensions as they have been raised in some contemporary writings of leading economists.

A CRISIS OF ECONOMIC FAITH

Long before the Protestant Reformation, there were many within the Roman Catholic Church who saw the failings of scholastic theology and preached the necessity of reform throughout the Church. Erasmus, a contemporary of Martin Luther, saw many of the same problems in the Catholic Church that Luther would condemn. However, as Tillich writes, Erasmus approached religious questions with an attitude of “detached

² See, for example, Robert H. Nelson, *Economics as Religion: From Samuelson to Chicago and Beyond* (University Park, PA: Penn State Press, 2001); and Robert H. Nelson, *Reaching for Heaven on Earth: The Theological Meaning of Economics* (Lanham, MD: Rowman & Littlefield, 1991).

analysis.” He was critical of the established church but it was “criticism ... of a rational kind, lacking in revolutionary aggressiveness.” The criticisms of Erasmus and earlier writers had largely failed to have much impact. By contrast, Luther “could not stand this skeptical attitude.” He “was a radical, in political as in other respects” (Tillich 1967, 238). It was Luther, of course, who eventually moved the world, overcoming a longstanding church inertia. The Pietists would later follow and, in some respects, move farther down the path originally blazed by Luther.

The economics profession today has its followers in the tradition of Erasmus. They see the problems of the profession and the need for reform, but approach this with an attitude of detachment and lack any real revolutionary commitment to change. Nevertheless, their criticisms, like those of Erasmus, are working to undermine the legitimacy of the established priesthood. Even many leading economists today no longer show much faith in the methods and character of contemporary economics. Indeed, William Davis recently surveyed economists about whether they believed in their own profession, and the findings showed a truly high degree of dissatisfaction and disillusionment (Davis 2004).

REFLECTIONS ON THE CONDITION OF ECONOMICS: *EJS* CENTENNIAL ISSUE

The internal crisis within economics at the end of the twentieth century was evident in a 1991 special issue of *The Economic Journal*. As a commemoration of its first hundred years of existence, the journal published a series of 22 articles on “The Next 100 Years” by leading economists. The articles provided an occasion for reflection on the record of the economics profession over the previous 100 years. Some of the articles were optimistic about the future of economics, even as they had significant criticisms to make. The great majority of the articles suggested that economics had become too narrow and that a widening of professional methods and a firmer empirical grounding for the discipline would be desirable.³

³ Besides the writers reviewed in this article, the centennial issue of *The Economic Journal* also included contributions by James Buchanan, John Kay, John Pencavel, Nicholas Stern, Joseph Stiglitz and Stephen Turnovsky. Agreeing in many respects with the authors whose critical views are examined in this article, these economists all expressed significant concerns for the current directions of the economics profession and stated the need for major

A considerable number of the articles, however, offered still stronger criticisms, showing a deep concern about scholastic tendencies that had emerged within the economics profession—occasionally even using the term “scholastic.” Few of these economists went so far as to advocate what I term pietism. Many of them suggested greater integration of economics with psychology, sociology, history, and other fields of study.

Here I review several of the articles and suggest that the complaints about topical scholasticism can be well understood as problems arising from a structural scholasticism that dominates contemporary economics.

Andrew Oswald

Dartmouth economist Andrew Oswald begins his contribution to the centennial forum of *The Economic Journal* by describing a sense of malaise concerning the directions of the profession.

Is Economics going in the right direction? Some people think not. Wassily Leontief has argued that our discipline has deteriorated into a second-rate branch of applied mathematics in which, unscientifically, researchers eschew empirical investigations. James Heckman says that the subject is “widely perceived to be discredited because it has so little empirical content and cares so little about developing it.” John Pencavel concludes that economists do not want applied work to be done, because it is likely to reveal the irrelevance of their hypotheses and undermine their ability to derive sweeping implications from theoretical models. (Oswald 1991, 75)

Oswald states that he is at least sympathetic to, if not in complete agreement with, the critics of the profession. Professional economics is in a “downward spiral” that reflects the influence of a “post-war generation of

changes. Less critical but still seeing the need for significant changes in the future practice of economics were Jagdish Bhagwati, Partha Dasgupta, and Richard Schmalensee. By contrast, the “true believers” in current economics among the centennial contributors were Peter Fishburn, Charles Plott, and Alvin Roth. Another contributor, J. Johnston, was also generally optimistic about the path of economics today. Finally, one contributor, Austin Robinson, had little to say about past or future economic methods; his article instead provides a survey of economic history.

mathematicians [who now] hold power” in the ranks of the profession. Among this group, “formal analytical ability,” as opposed to empirical and other more useful knowledge, “is the criterion for advancement.” Furthermore, young economists are being molded to fit this model because, “believing themselves to be an elite, the ruling class [in economics] aim to create future generations in their own image.” Their manner of exercising control over the activities of the economics profession is “by accepting for publication only certain kinds of articles, by recommending for promotion young mathematical economists, and by changing graduate courses to stress technical skills at which they excel” (Oswald 1991, 75). The medieval priesthood as well maintained its scholastic orthodoxies by asserting rigorous controls over admission to its ranks.

Investigations by Oswald and others showed that about half of the articles published in leading journals by economists had no data. The economics profession sees itself as following in the path of physics and chemistry, but only 12 percent of physics articles and almost none of the articles in chemistry had no data (Oswald 1991, 75). The difference might be explained partly by the fact that the subject matter of economics is much broader than physics or chemistry and appropriate data simply do not exist. Oswald concludes, however, that much of the work of the economics profession is a disguised “kind of mathematical philosophy,” engaging in abstract reasoning of doubtful utility, a trend he finds “hard to believe . . . a desirable state of affairs” (Oswald 1991, 78).

Oswald’s description sounds much like the criticisms of scholastic writings at the low points from which the negative connotations of “scholastic” are derived (the high points of the scholastics included the writings of Thomas Aquinas and others among the great medieval philosophers). As the intellectual historian John Herman Randall comments, there came a time when “the medieval intellect had . . . built as large an edifice as it could hope to with the materials at hand, and had commenced those fine drawn distinctions which have given its debased form so evil a name” (Randall 1926, 213). There was in such scholastic thought a “preoccupation with mere forms of knowledge” that worked well for “expounding and developing an authoritative body of principles.” However, “it could teach little that was not already known” (Randall 1926, 214). As Francis Bacon complained, the scholastic manner of argument “brings forth indeed cobwebs of learning, admirable for the fineness of thread and work, but of no substance or profit” (quoted in Randall 1926, 214). Nevertheless, it did offer, as Randall states, “a means of increasing man’s power over his fellow man, enhancing his reputation and his purse by

victory in disputation” (Randall 1926, 214). It all sounds very much like the current economics profession, as described by many of the writers in *The Economic Journal*.

William Baumol

Writing in this same issue, some of the best known names in economics—members of the profession whose reputations were partly based on technical forms of exposition—show deep concern about scholastic tendencies within the field. William Baumol finds that a “peril” facing economics is that “few specialized students are allowed to proceed without devoting a very considerable portion of their time to the acquisition of mathematical tools, and they often come away feeling that any piece of writing they produce will automatically be rejected as unworthy if it is not liberally sprinkled with an array of algebraic symbols.” If they engage in “the pursuit of alternative approaches,” their work will “not [be] respected” by the leadership of the economics profession (Baumol 1991, 2). In looking towards the next century, Baumol states that “it should by now be obvious that I am hoping that the future will bring some decrease in the display of technique for its own sake, with models constructed so as to increase what they tell us about the workings of the economy rather than just displaying the properties of some analytical procedure” (Baumol 1991, 6). The church of economics, as Baumol is saying, is straying farther and farther from the real world into elaborate displays of mere technical competence in mathematical reasoning.

Milton Friedman

Milton Friedman is generally supportive of the turn during the twentieth century of the economics profession towards greater use of mathematical and statistical methods. However, much like Baumol, he finds that things have gone too far. Indeed, Friedman declares that the “reliance on mathematics and econometrics” has reached “the point of vanishing returns.” The use of mathematics is no longer making a contribution to economic understanding, but has become an end in itself. As Friedman comments, “again and again, I have read articles written primarily in mathematics, in which the central conclusions and reasoning could readily have been restated in English” (Friedman 1991, 36). One of the

revolutionary changes of the Protestant Reformation was to translate the Bible and other religious writings from Greek and Latin into the language of the ordinary people. Among the reasons for Luther's great fame, he was the first to translate the Bible into ordinary German. Friedman, it would seem, sees a similar need in economics for translations, when there are worthwhile ideas, into the common language of today.

Friedman finds that, even in 1930, under the editorship of John Maynard Keynes, the entire volume 40 of *The Economic Journal* contained one page that included mathematical symbols. But mathematics after World War II became the language by which economists in the twentieth century sought to assert their priestly prerogatives. Unfortunately, as Friedman concludes, this has not led to corresponding increases in economic understanding. "[T]o summarize," he writes, "there has been little change in the major issues occupying the attention of economists: they are very much the same as those that Adam Smith dealt with more than two centuries ago. Moreover, there has not been a major sea change in our understanding of these issues." In physics and chemistry the writings of 200 years ago are a mere historical curiosity. But it is still possible to "read the *Wealth of Nations* and David Hume's essays *Of Money* and *Of Interest* with pleasure and intellectual profit" (Friedman 1991, 37). In re-examining old volumes of *The Economic Journal*, Friedman is struck by how "the substance of professional economic discussion has remained remarkably unchanged over the past century" since the first volume was published. If the substance was not much different, to be sure, "the language" has changed "drastically" (Friedman 1991, 33). Displays of virtuosity in the new language of mathematics have become more important for many economists than the development of real economic enlightenment.

Friedman also is pessimistic in seeing little gain in the quality of economic understanding from the nineteenth to the twentieth centuries; it is poor in both periods. He quotes a statement of an economist W.J. Ashley in 1907 that "when one looks back on a century of economic teaching and writing, the chief lesson should, I feel, be one of caution and modesty, and especially when we approach the burning issues of our own day. We economists . . . have been so often in the wrong!" Friedman declares that this conclusion from 1907 "can serve as mine in 1990" (Friedman 1991, 39). The great commitment to formal quantitative rigor in economic methods of the twentieth century, as Friedman concludes, has done little to improve economic judgments.

Michio Morishima

Other economists who contributed to *The Economic Journal* reviewed the status of general equilibrium theory (GET), once considered the highest grounds of theoretical development within the economics profession. Despite its great prestige, the reviewers find evidence of the spread of scholastic tendencies in the work of general equilibrium economists. GET theorists, Michio Morishima finds, have “sunk into excessive mental aestheticism” (Morishima 1991, 70). Economists poured great resources into analysis of “the world of GET [which] is in fact a dream world” (Morishima 1991, 71). Yet, it is possible to achieve the highest levels of professional prestige in economics by developing the full contours of this fantasy story. There are many economists who “expend their energies on competing with each other in demonstrations of intellectual and theoretical ability.” For those who are successful in this manner, they often “regard as their inferiors those who contend [that there exists] the need to observe the real world.” For Morishima, however, all this is “a palpable symptom of scientific degeneration” within the community of high economic theory (Morishima 1991, 74).

Frank Hahn

Frank Hahn, a leading economist who himself gained fame as a general equilibrium theorist, is equally skeptical. He predicts the “demise” of GET, although he does not necessarily agree with the view of others that in every case “pure theory is scholastic and so by implication bound to be irrelevant to the world.” The method of pure theory involves “the activity of deducing implications from a small number of fundamental axioms.” Despite its current problems, Hahn thinks that this effort in the past often led “to beauty and to surprise.” The work of general equilibrium theorists, Hahn suggests, has been “crucial to our understanding of decentralized economies” (Hahn 1991, 47).

The problem for Hahn is that general equilibrium theory now will be much less useful in addressing “the next crucial questions” for economics. Indeed, “almost none of them,” Hahn declares, “can be answered by the old procedures.” In future economic research, it will be necessary to introduce “psychological, sociological, and historical postulates.” The

“axiom of rationality” will have to be significantly relaxed. Economics will have to become “a ‘softer’ subject than it is now” (Hahn 1991, 47). There will have to be an embrace of the concerns and methods of the fields of “history and sociology and biology” (Hahn 1991, 50).

The language of mathematics, as Hahn believes, will probably be less helpful in any such eclectic effort to integrate so many fields. Indeed, ordinary English might be required to bring together the insights from a wide range of scholarly sources. Professional credentials in any one specialized body of expertise might count for little; any intelligent person might be as qualified as any other person in synthesizing diverse areas of knowledge. It might require a modern “Economic Reformation” in which the voices of the professional priesthoods no longer carry special authority.

So far, however, the economics profession is resisting the transition to any such new world. Hahn sees this resistance as characteristic of a religion under challenge: “one often encounters increased orthodoxy among some just when religion is on the decline.” For the current economics profession, “it is clear that this sort of thing heralds the decadence of endeavor just as clearly as Trajan’s column heralded the decadence of Rome. It is the last twitch and gasp of a dying method.” The scholastic thinkers of a later Christian era in Rome had also been swept aside in the end by the tides of history. By continuing to focus on “the maximization of a representative agent’s utility over an infinite future,” the result is that economists are “ignoring every one of the questions now pressing for attention” from the profession (Hahn 1991, 49). They might well also soon end up as losers in the intellectual tides of their own times (see Nelson 2001).

Edmund Malinvaud

Another leading technical economist, Edmund Malinvaud, offered yet another pessimistic view. In assessing economics since World War II, Malinvaud declares that these years “were obviously marked first by a wave of optimism, then by the painful realization that most of the initial beliefs were the product of delusion. This applies whether one considers the broad development issues or the more modest current problems of industrial countries.” In the 1950s large numbers of economists believed that their work would “lead to international economic order; it will gear development in the Third World; it will show the way to good socio-economic performance in alternative systems to capitalism.” As the events of the

second half of the twentieth century unfolded, however, these optimistic beliefs were not realized. “[T]he beliefs appear to have been mainly unwarranted, following from wishful thinking and from bold or loose extrapolations of what economics really knew” (Malinvaud 1991, 65).

In the later decades of the twentieth century, economists shifted their optimism towards the management of the market system. Yet, as Malinvaud assesses this more recent history, “the same sequence of confidence and disappointment occurred with respect to the role of economic management in market economies, whether it concerned allocation of resources, distribution of welfare or macroeconomic stabilisation.” Part of the problem has been the failure of economists to understand that “public management is never a purely economic matter and cannot be immune from political interference, if only because the notion of an objective to be achieved can seldom be precisely defined beforehand.” Another problem is that “side effects that had been taken as negligible turned out to be determinant.” On the whole, there has been a demonstrated “inability to solve the real problems” that has acted to undermine the earlier high hopes of the profession (Malinvaud 1991, 65).

Despite the failures of the past, Malinvaud is optimistic that a new and better economics will emerge in the future. Like Hahn, he believes that this will require the introduction of new assumptions and ways of thinking from other fields of study. Economists will have to incorporate the insights of “psychologists, sociologists, political sciences” and other areas into their work. There will be a need for “supplementary information on physical, technological, institutional, or social constraints.” It will be necessary to “recognize the limits of the dominant concepts of economic rationality and economic equilibrium” (Malinvaud 1991, 68). The result might look more like the old subjects of political economy or moral philosophy, rather than the current efforts of economists to emulate the physical sciences. “[A]fter a period of doubts the usefulness of economics as a normative science will again be recognized” (Malinvaud 1991, 67).

To be sure, it will depend on the future actions of economists themselves. Malinvaud acknowledges the risk that “seriously exists that the discipline progressively loses touch with real problems, develops on its own into a scholastic [exercise] and becomes less and less significant for layman’s concerns.” Indeed, there are warning “signs of such an evolution” in the current activities of economists. There are “great efforts ... being spent [at present] for solving problems whose ultimate relevance can only be very indirect” (Malinvaud 1991, 66).

Interlude: How to Avoid Topical Scholasticism

Admittedly, even if members of the economics profession accept the necessity to widen the areas and the methods of economic inquiry, it might not mean the end of priestly prerogatives. Economists might not use as much mathematics, but the use of other professional jargon might still be required in order to be taken seriously within the ranks of the profession. In other areas of professional activity such as the law, this verbal method of defending professional “turf” has long been employed. Nevertheless, if economists are expected to write in plain language, and to incorporate the ideas of various fields of scholarship into their work, it will be more difficult to assert a professional monopoly on economic knowledge. Interested members of the general public will have a greater ability to assess the relevance of what economists write and the merits of their arguments. Thinkers outside the economics profession will be able to compete on a more equal basis.

If such a “reformation” of economics takes place, it may well turn in the direction of a new pietism. As one theologian comments, the followers of Pietism in the late seventeenth century, and for much of the eighteenth century, were “reacting against an elite, remote, self-satisfied professionalism” among the established clergy. In the small communities in which they gathered, the “priestly functions were [instead] practiced by laity.” The theology of Pietism “was oriented toward the practical implementation of behavioural change rather than its theoretical aspects”—as one might say today, towards real world policy argumentation rather than formalistic exercises. A main feature of “pietistic faith and practice” was that it “exhibited a broad spirit of tolerance and a primitive ecumenism” (Oden 1972, 80). Pietists did not impose strong theological preconditions for participation in their community discussions; they were willing to draw from many religious traditions in their intense search for a proper individual relationship with God. Economists, as many contributors to *The Economic Journal* were suggesting, might similarly have to widen their sources of information and thinking in the development of their own economic ideas.

Jack Wiseman

Most of the contributors to *The Economic Journal* were leading economists whose work fell within the mainstream of economic research. The *Journal*, however, opened its pages to two longstanding economic critics. As Jack Wiseman noted, “I have long preached that mainstream economics is fundamentally flawed.” Although his message had not yet been accepted, Wiseman (who died soon after in 1991) was confident that time was on his side. He believed that “the heretics grow in numbers” and he was “increasingly confident that they will be tomorrow’s priests” (Wiseman 1991, 149).

As Wiseman saw trends within economics in 1990, “the need for a new paradigm is coming to be more generally accepted.” There is “growing dissatisfaction with the dominant neo-classical orthodoxy” (Wiseman 1991, 150). Among a group of “evolutionary economists,” for example, they “dismiss the behavioural assumptions of neoclassical economics as destructively simplistic.” There were many “dissident groups,” also including the Austrian school, the new institutionalist school, the public choice school, behavioural economics, and the “radical subjectivists personified by Shackle.” For Wiseman, the various criticisms of the economic mainstream demonstrated that “adaptation” would not be enough: it would be necessary to have a “fundamental reappraisal” of the work of professional economists (Wiseman 1991, 151). As it seems, the great need today is for a new Luther or Calvin of economics.

John Kenneth Galbraith

Only one economist who was asked to contribute to *The Economic Journal* might have been so bold as to entertain any such personal ambitions. Since the 1950s, John Kenneth Galbraith has been writing with great success for audiences of mostly non-economists. He has been an economic heretic not only in this respect, but in the many fierce attacks he has long directed at the work of mainstream economists.⁴ He continues in that vein in his *Economic Journal* contribution, declaring that current members of the profession devote their main efforts to scholastic work that “allows of an infinity of technical refinement within an unchanging context.” Economics

⁴ See for example, Galbraith’s *Economics in Perspective: A Critical History* (Boston: Houghton Mifflin, 1987).

has developed an ingrown culture in which the successful feel “a rewarding sense of superiority as compared with those who have not similarly penetrated the complexities.” Some of these economists are so absorbed in their own imaginary creations that they fail even to recognize the drastic departures from the real world economic circumstances. Yet, in looking to the longer run, as Galbraith believes, it is inevitable that they will face “intellectual obsolescence” and “increasing irrelevance” (Galbraith 1991, 41). It has not happened yet, however. Galbraith has long advocated a revolution in economic methods, but his personal crusade has made little impact on the economic mainstream.

PIETISM AS BOTTOM-UP ENLIGHTENMENT

A pietistic approach to the search for knowledge offers no guarantees of the final economic and policy outcomes. There are “left” economists such as Galbraith and libertarian economists such as Henry Hazlitt who have written in plain language and whose main influence has been felt outside the mainstream of professional economics. The Protestant Reformation earlier had opened the discussions of theology to a much wider range of participants, thus giving a new life to religion. Protestantism did not, however, yield religious closure. To this day, Protestant denominations compete fiercely with one another for followers—a “free market of religion” that replaced the tight control of the Catholic Church on the acceptable bounds of religious expression. A pietistic approach to economic knowledge today would similarly offer a “free market in economic ideas” in which the efforts of professional economists to assert powerful priestly privileges would be significantly curtailed. Getting down to policy brass tacks, as Protestantism long ago abolished much of the edifice of clerical privilege, this might mean ending academic welfare as dispensed today in support of the priestly hierarchy and in other government subsidies to the scholastic apparatus.

To be sure, the history of the Protestant Reformation illustrates the potential hazards of the pietistic approach as well. Lacking central authority, intellectual confusion may result. The Protestant Reformation led to religious warfare covering much of Europe for more than a century—at great cost in lives and property. There is a balance required between the scholastic and the pietistic approaches to the search for religious and

economic truth. In professional economics today, however, the scholastic mode has become, in my judgment, much too powerful. It will be necessary to find a better compromise by turning towards a more pietistic approach to economic learning.

Robert Solow: Champion of the Existing Structure

About the same time that the contributors to *The Economic Journal* were offering their commentaries on the state of economics, another group of economists were reflecting on the economics profession in an edited book collection, *The Spread of Economic Ideas* (1989). The writings there were often consistent with the many pessimistic views expressed in *The Economic Journal*. For example, the editors, A.W. Coats and David Colander, note in the introductory essay that the economics profession engages in many internal controversies “at times resembling medieval theological disputations” (Coats and Colander 1989, 4). In a later chapter, David Colander suggests with respect to the research efforts of the profession that “the emperor has no clothes” (Colander 1989, 36).

However, one contributing author was Robert Solow, the MIT economist and winner of the Nobel prize in economics in 1987. Solow’s views departed significantly from those described above. As Solow notes, most of the other chapters in the book reflect “a sense that our profession is marked by utter confusion and loss of confidence and bearings.” However, “I do not share that feeling, not at all, nor do most of the functioning economists I hang out with” (Solow 1989a, 37). Solow does lament that economic knowledge is often ignored in the political process. However, this is not due to the failings of economics. Overall, Solow portrays, instead, a world of economic priests whose valid economic truths are simply unwelcome and unheard in a sinful world.

Solow believes that economics, properly done, can and should be value-neutral—even though many would question that this is a possibility, even in concept. As he states, it is true that “the positive and the moral aspects of economics are very much intertwined. But honesty and clarity require that in talking about economics we try our hardest to separate them” (Solow 1989a, 38). Solow has a clear model in his mind of how the world of economic policy making should work. First, economic scientists will commit themselves to objective research to discover the theoretical structures that underlie the workings of an economic system. Many economists will try, but most of the deepest thinking will be produced by a

select few—those who are the best trained, and have the highest scientific skills, and are mostly located in the leading universities.

In the next step, their discoveries will be absorbed within the economics profession itself. From there, the ideas somehow must trickle down and filter into the public mind and the political process. As Solow describes this overall process, “economic ideas—ranging from new, but unelaborated concepts, through isolated propositions about causality, all the way to full-blown theories—arise in the highbrow part of the economics profession and then diffuse, first within the profession, and then sometimes outside it to journalists, bureaucrats, politicians, and other citizens” (Solow 1989b, 75). Religious—and now economic—truth originates in the highest levels of the church, is then communicated to the wider priesthood, and finally reaches the laity.

In the medieval development of scholastic theology, and now in the economics profession, the ordinary person has very little to contribute to improved knowledge. Solow’s vision is strictly top-down. He says that “there is also a backflow of ideas from the World Out There to the economics profession, but those are hardly ever economic ideas. They are rather social beliefs, priorities, or ideological conceptions” (Solow 1989b, 75). The public ideas, moreover, often unfortunately influence the reception to the scientific knowledge being offered by economic theorists.

Here, Solow confesses to his own feelings of pessimism. The public distorts what it hears from economists. “[O]ne’s fear, of course, is that this distortion [of the economic message] is often substantial and systematic” (Solow 1989b, 75). The result is that the public understanding of economic theories may “bear little resemblance to the original ideas, the ‘real’ ideas.” It may even mean that “what is finally transmitted is merely false” (Solow 1989b, 76). Solow offers an example from the development of the 1986 tax law, one of the most important pieces of U.S. economic legislation of the 1980s. In listening at the time to an influential member of the Congress explain this law, Solow finds that the Congressman is speaking “meaningless drivel.” What had begun as a valid economic principle was now only the “residue of economic ideas,” and in the thinking of a leading Congressman amounted to “nonsense” (Solow 1989b, 79). The application of economic science could make for a far better world, but all too often the higher learning that economic experts have to offer is ignored by the politicians.

Thus, it turns out that, while Solow largely exempts the economics profession itself from the blame, he shares the sense of confusion and uncertainty of purpose felt by many of his colleagues concerning the policy

role of the economics profession. How can valid technical economic concepts be put to practical use, when the political leadership is typically incapable of understanding these concepts? As Solow laments, “good economics is bound to be complicated. Good economics is also bound to be uncertain” (Solow 1989b, 82). One option would be for the politicians simply to trust the economists—as a patient trusts a heart surgeon. At one time, that was the hope of the economics profession, but it was not to be.

The political leaders who are the intended audience for economic knowledge, as Solow finds, are “rarely interested in narrow economic policy for its own sake.” They are unwilling to listen to “the complexities and uncertainties of economic analysis.” The politicians want simple answers. Some economists, anxious to be heard and to please their audience, are in fact willing to bend to political pressures. As Solow laments, in such cases “the result is that we [economists] pretend to answer questions far beyond the capacity of our observational material to provide credible and reliable evidence” (Solow 1989b, 82).

I find much truth in Solow’s reflections, of course. Indeed, the problems Solow notes are a main reason why I generally favor a much smaller government, especially at the national level in the United States. But still, Solow sounds somewhat like the Pope in the Vatican confronting a world where abortion, sexual license, and other sinfulness reigns triumphant. The high priests of economics have the knowledge to improve the world but few will listen. Heresy reigns over much of the earth. Rather than compromise with evil, Solow suggests that perhaps it will be necessary to withdraw to the monastery. If their theories must be complicated, and yet the world demands simple answers, he says that perhaps the best response is “a little more silence” among economists (Solow 1989b, 83).

CONCLUSION: TOWARD A MORE PIETISTIC ECONOMICS

Economists such as Solow implicitly held to a model of the political world for much of the twentieth century. They disavowed being “political” yet had a strong normative vision of the proper workings of the political system. In this political theory, economists produce objective economic knowledge. Politicians make the value judgments for society in separate arenas and then seek the definitive advice of economists in

realizing these values and the associated social outcomes. It all assumes that there can be a separation of the political process into two distinct domains—one of scientific expertise and the other of the determination of social values.

However, this model was always doubtful and few subscribe to it any more, including most economists. If the world had worked according to the political ideals professed by past economists, the result would have been, in practice, to turn over the making of government policy to the economists themselves (which is probably part of the model's appeal to many of those economists). The problem is that it is typically impossible to separate value decisions and technical decisions. Indeed, the very distinction between fact and value is murky at best: Important realms of fact concern the values that people in fact cherish. In the real world, "social values" typically emerge incrementally, the outcome of a large number of administrative and other small decisions. If economists and other experts are given significant control over these decisions, they will end up, in effect, determining much of the direction of social values. Economists will truly be the moral priesthood of society.

Economists like Solow should reconsider the political model that has long existed in the back of their minds. They should reconsider the manner of doing economics. Economics that is not widely understandable, whose insights cannot be incorporated directly into plain language that is accessible to public discourse, may be close to useless for public purposes. If economists continue to produce work that has such limited public utility, the political process may turn against the funding of economists with tax-dollars. In the longer run the support is not likely to continue indefinitely for a large commitment of resources that is directed mainly to the satisfaction of the aesthetic tastes of economic model builders and statisticians.

In terms of actual economic understanding, the costs of a more public-discourse orientation are likely to be minimal. Indeed, a new model of economics in plain language may well improve the quality of future economic knowledge. Economists in the second half of the twentieth century turned away from wider empirical inquiries to explore the logical deductions from a small set of "simplifying" assumptions and whether the world actually conformed to the resulting abstract predictions. However, many important influences on the economic system do not lend themselves to such a simplified mathematical expression. Rather than adopt less formal methods of inquiry, economists often chose simply to ignore these influences.

When outside critics complained that many important economic factors were being left out of the analysis, the economics profession dismissed these critics as lacking professional standing. Like any group that is permitted to exclude competitors, an isolated and autonomous economics profession often produced materials of little use or interest to the rest of the world. In a word, they had become another generation of scholastics who mistook the applause of fellow scholastics for real accomplishment.

These tendencies within the contemporary academy admittedly are not limited to the members of the economics profession. Much as Adam Smith said of the wider university world more than 200 years ago, the intellectual historian Alan Kors now observes of our own time that:

I don't think we yet know what would happen if, in addition to standing up within our own institutions, we made the broader public aware of how many areas of the social sciences and the humanities have become intellectually marginalized, and of how much better educated their children would be from reading things Left, Right and Center produced by think tanks rather than from studying the social sciences at a university In the social sciences, and in many of the humanities, the most interesting things are occurring outside the university. In some ways it reminds me of the relationship of the universities to the physical sciences in the seventeenth and eighteenth centuries. There was an ideological commitment to a certain Aristotelian scholasticism in the seventeenth century that forced the new experimental scientists to find homes outside the universities. As a result, the most interesting science in the seventeenth century was not done at the University of Paris or at Oxford, but in the Royal Society or the Academy of Sciences in France, or in diverse private societies throughout western Germany, northern Italy, and, indeed, France and England That's the case now in the social sciences. What happens in the American Sociological Association is trivial. But what's coming out of certain think tanks and certain foundations and certain institutes is very exciting and much more central to the real debates about the problems of American society. (Kors 1988, 88-89)

Whether the result of internal forces within the economics profession or external pressures, the scholastic economics of the present cries out for a new reformation. It would require a turn in economics in a pietist direction. In the Protestant churches of Pietism, the leadership was required to speak the language of the laity. The old distinction between an authoritative priesthood and the ordinary people was abolished.

Ordinary people in the future may also have to be able to understand economic arguments, as Protestants once argued that they must study the bible for themselves. As theologian Thomas Oden writes, "Protestant pietism [was] . . . profoundly instrumental in liberating the western tradition from the strictures of scholastic orthodoxy and in helping to introduce . . . [this tradition] to the modern world" (Oden 1972, 66). The economics of the future will have to incorporate, as one might say, a new economic pietism that will help to liberate the twenty-first century from the scholasticism that came to dominate the practice of professional economics over the course of the twentieth century.

REFERENCES

- Baumol, William J.** 1991. Toward a Newer Economics: The Future Lies Ahead! *Economic Journal* 101(January): 1-8.
- Coats A. W. and Collander, David C.** eds. 1989. An Introduction to the Spread of Economic Ideas. In *The Spread of Economic Ideas*. ed. A. W. Coats and David C. Colander. New York: Cambridge University Press, 1-19.
- Colander, David C.** 1989. The Invisible Hand of Truth. In *The Spread of Economic Ideas*. ed. A. W. Coats and David C. Colander. New York: Cambridge University Press, 31-36.
- Cragg, Gerald R.** 1970. *The Church and the Age of Reason, 1648-1789*. New York: Penguin Books.
- Davis, William.** 2004. Preference Falsification in the Economics Profession. *Econ Journal Watch* 1(2): 359-398.
- Friedman, Milton.** 1991. Old Wine in New Bottles. *Economic Journal* 101(January): 33-40.

- Galbraith, John Kenneth.** 1987. *Economics in Perspective: A Critical History*. Boston: Houghton Mifflin.
- _____. 1991. Economics in the Century Ahead. *Economic Journal* 101(January): 41-46.
- Hahn, Frank.** 1991. The Next Hundred Years. *Economic Journal* 101(January): 47-50.
- Kors, Alan C.** 1988. Voces Academicæ: Thought Reform and Education: A View from the University of Pennsylvania. Professor Alan C. Kors, interviewed by Carol Iannone. *Academic Questions* 1(4): 75-89.
- Malinvaud, E.** 1991. The Next Fifty Years. *Economic Journal*. 101(January): 64-68.
- Morishima, Michio.** 1991. General Equilibrium Theory in the Twenty-First Century. *Economic Journal* 101(January): 69-74.
- Oden, Thomas C.** 1972. *The Intensive Group Experience: The New Pietism*. Philadelphia: Westminster Press.
- Oswald, Andrew.** 1991. Progress and Microeconomic Data. *Economic Journal* 101(January): 75-80.
- Randall, John Herman.** 1976 [1926]. *The Making of the Modern Mind: A Survey of the Intellectual Background of the Present Age*. New York: Columbia University Press.
- Smith, Adam.** 1937 [1776]. *The Wealth of Nations*. New York: Modern Library.
- Solow, Robert M.** 1989a. Faith, Hope, and Clarity. In *The Spread of Economic Ideas*. Ed. A.W. Coats and David C. Colander. New York: Cambridge University Press, 37-41.
- Solow, Robert M.** 1989b. How Economic Ideas Turn To Mush. In *The Spread of Economic Ideas*. Ed. A.W. Coats and David C. Colander. New York: Cambridge University Press, 75-83.
- Tillich, Paul.** 1967. *A History of Christian Thought: From Its Judaic and Hellenistic Origins to Existentialism*. New York: Simon and Schuster.
- Wiseman, Jack.** 1991. The Black Box. *Economic Journal* 101(January): 149-155.

ABOUT THE AUTHOR



Robert H. Nelson is a professor at the School of Public Policy of the University of Maryland. He is the author of six books, including *Zoning and Property Rights* (MIT Press, 1977); *The Making of Federal Coal Policy* (Duke University Press, 1983); *A Burning Issue: A Case for Abolishing the U.S. Forest Service* (Rowman & Littlefield, 2000); and *Economics as Religion: From Samuelson to Chicago and Beyond* (Penn State Press, 2001). Mr. Nelson has also written for the *Washington Post*, *Wall Street Journal*, *Los Angeles Times*, *The Weekly Standard*, *The Financial Times*, *Reason*, *Technology Review*, *Environment*, and many other popular publications. He was a columnist for *Forbes* magazine from 1993 to 2000. He worked in the Office of Policy Analysis of the Interior Department—the principal policy office serving the Secretary of the Interior—from 1975 to 1993. His email address is: nelsonr@umd.edu.

INVESTIGATING THE APPARATUS

The National Research Council Ranking of Research Universities: Its Impact on Research in Economics

RANDALL G. HOLCOMBE*

[Abstract, Keywords, JEL Codes](#)

THE ADMINISTRATORS AND FACULTY AT MANY UNIVERSITIES view their missions as teaching, research, and service, and among a substantial subset, research is the central mission upon which the other two missions are built. Administrators and faculty at research universities view their institutions as more prestigious than universities that focus primarily on teaching, and among research institutions there is also a hierarchy based on the quality of the faculty and the quantity and quality of the research done at the institution. Most people could recite a list of the most prestigious research universities—Chicago, Harvard, MIT, Stanford, and so forth—but quality is difficult to quantify, and people may have a harder time evaluating universities farther down the quality scale without additional information. Nevertheless, a measure of quality might be desirable for many reasons. Students choosing a graduate school could benefit from such an indicator, for example, and university administrators might be interested in some measure of how departments in their institutions compare with those in other universities. In response to the desire for some quantitative ranking of research universities, the National Research Council (NRC) has produced

* Department of Economics, Florida State University.

several studies ranking graduate programs. This paper discusses some of the effects of the NRC rankings on research activities in economics departments.

The NRC rankings quantify the quality of research done in economics departments, but they also can affect the type of research done in economics departments. The NRC rankings affect research activity within some departments because the administrations at some universities have set an explicit goal of improving their NRC rankings. For example, at my institution, Florida State University, faculty have been told that one of our goals is to improve our NRC ranking, the administration has communicated to faculty the details regarding how the NRC ranks faculty, and individual faculty are asked to report their own research activities following the NRC format. This way, each faculty member is able to judge his or her own contribution to the department's NRC ranking.

One result is that faculty members can be compared with each other based on the NRC criteria, and given the administration's goals, this affects the department's evaluation of faculty research, decisions on hiring, promotion, and tenure, and sends a signal to faculty regarding how their activities contribute to the overall institutional goal of improving the department's NRC ranking.¹ Activities that are counted by the NRC get increased emphasis, while those the NRC does not count are deemphasized. I have discussed this phenomenon with colleagues at other mid-level universities (Florida State's economics department ranked 60th in the most recent NRC study) and found that their institutions have similar goals and are using similar methods.² Although it does not appear that many other universities are adopting NRC methods as directly as FSU is, Mr. James Voytuk, who aided the production of the 1995 NRC study, and who is currently the chair of NRC's Committee for the Study of Research-Doctorate Programs in the United States, reports that he observes strong interest in NRC's rating methods from representatives of the Council of Graduate Schools and similar organizations.³ Before considering the effects

¹ Hires right out of graduate school may have little to show in any of these dimensions, but senior hires will be evaluated this way and junior hires will be evaluated on their potential for scoring high using these criteria.

² Klein and Chiang (2004a) present survey results indicating that citation counts are used more at mid-level universities than at the top schools. Their survey obtained responses from a subset of universities and I have discussed these issues with colleagues at other universities that were not among the respondents to Klein and Chiang's survey.

³ Mr. Voytuk conveyed this observation in a telephone conversation with Daniel Klein on 21 September 2004.

of this competition in more detail, the NRC ranking procedure will be described and discussed.

THE NRC RANKING PROCEDURE

The NRC is an outgrowth of the National Academy of Sciences (NAS), which was created in 1863 by an Act of Congress to promote science and technology.⁴ The Civil War prompted the NAS's creation, and it investigated issues like how to protect the bottoms of iron ships, and how to correct for magnetic compass deviation. The NRC was similarly founded by an Act of Congress in 1916 to coordinate scientific and technological research and development, this time related to the outbreak of World War I. The NAS promotes and funds research projects and the NRC, as a part of NAS, has undertaken several studies evaluating the quality of academic departments at research universities.

The results of the most recent NRC study are found in Goldberger, Maher, and Flattau (1995), cited as NRC throughout the remainder of this essay, which ranks 3,634 doctoral programs in 41 fields at 274 universities in the United States (NRC, 19). Of 135 programs that awarded doctoral degrees in economics from 1986 to 1992, a total of 106 economics programs are evaluated (NRC, 20). The rankings were done by mailing questionnaires to faculty at those institutions and asking them to evaluate a subset of other programs in their field. Respondents were asked to rank the scholarly quality of program faculty (as distinguished, strong, good, adequate, marginal, or not sufficient for doctoral education) and the effectiveness of the program in educating research scholars and scientists (as extremely effective, reasonably effective, minimally effective, or not effective). The responses were then placed on a numeric scale and the rankings aggregated to give each program a score of 1-5 in each of the two areas: quality of faculty, and effectiveness of the program. Departments are listed in the NRC volume in rank order determined by the subjective evaluation of the quality of their faculty, and because of this, the subjective quality-of-faculty measure is viewed as a department's overall NRC ranking.

⁴ Background information on the NAS and NRC can be found at www.nationalacademies.org/about.

The NRC study ranks individual doctoral programs at universities, but it does not rank the universities.⁵

The NRC notes that there are several objective measures that are highly correlated with the subjective ratings of departments, and reports three in their results: publications per faculty member, citations to faculty publications, and research grants to faculty. If the rankings were given in a different format—for example, if departments were listed in alphabetical order—readers might be more inclined to weigh different metrics when evaluating a department. For example, if one department ranked 12th in citations and 15th in the subjective evaluation of faculty quality, that department might (or might not) be viewed as a better department than one that ranked 15th citations and 12th in the subjective evaluation of faculty quality. Because the rankings are presented in order of the subjective faculty quality rankings, however, it would appear to most readers that the first department ranks 15th overall while the second ranks 12th. The way the rankings are presented has a big impact on the way they are perceived. The rating for effectiveness in educating students is in the study right beside the rating for quality of faculty, but because schools are ranked by the survey results on quality of faculty, that metric is clearly the one that counts most. The tables give a total of 20 columns of data, including faculty publications, citations, grants, mean year to get a degree, and other factors that might be important indicators of quality. Yet by the table's design, ultimately it is the survey results on faculty quality that determines the quality-ranking of a department, and the other information is there if people want it.

Nearly 600 pages of the 740 page NRC study are tables that examine and present the rankings of doctoral programs. Each program has 20 data fields describing it, including the current rankings and change in ranking from the earlier study, demographic information such as number of faculty, students, female students, minority students, US students, and faculty publications, citations, and grants. While the departments are presented in rank-order of their faculty quality as measured by the survey, the data are available in electronic form through the NRC, and the study notes that “many other types of analyses are possible and encouraged” (NRC, 15 and 59). Discussions with colleagues at other universities have revealed that many universities in addition to my own have taken this suggestion

⁵ Rankings of doctoral programs based on faculty opinion goes back at least to 1925 (NRC, 10), and the NRC first produced such a ranking in 1982 (Jones, Lindzey, and Coggeshall, 1982). The purpose of the 1995 NRC study was to update and improve on what was presented in the 1982 study.

seriously and used the NRC data to assess the quality of their departments, and to look for ways to improve their NRC rankings. One result of this is that faculty evaluations place more emphasis on those indicators reported by the NRC. Apparently, the idea is that if a department improves in those objectively-measurable areas, their subjective survey rankings will also tend to rise.

QUANTIFYING QUALITY

One can see the justifications for trying to quantify the quality of a program, such as giving potential students a rough indication of the relative qualities of graduate programs they might be considering. But for purposes such as this, it should be apparent that quality is not a single-dimensioned characteristic. Graduate programs have different strengths, and a program well-suited for one student may not meet the interests of others. While people might recognize that some programs are clearly higher-quality than others, for some purposes there will be no way that, in an objective sense, programs can be strictly rank-ordered by quality. Yet by creating this type of ranking, a metric is created that enables university administrators to compare their departments with others and to try to improve their quality, as measured by that ranking.⁶

The NRC (3, 40, and 422) notes the correlation between the quality ranking and the objective indicators of publications, citations, and grants, so it is a short step for university administrators to conclude that their departments could increase their NRC rankings by improving in these dimensions. Therefore, one result of the publication of the NRC study is that university administrators are emphasizing those metrics as performance measures of their faculty. Thus, it is reasonable to examine those criteria and see what incentives they imply for economic research.

⁶ See Toutkoushian et al. (2003) for a justification of research rankings and a methodology for calculating them from the ISI publication database.

Journal Publications

While the NRC study discusses faculty publications, what it actually measures is publications in journals that are compiled by the Institute for Scientific Information (ISI). The ISI maintains a database that includes information from a subset of all academic journals, including 163 economics journals, and only articles in those journals are counted in the NRC publication methodology. The number of economics journals covered is substantial, but many economics journals are not included in the list, and the list does not include books or monographs.⁷ This creates an obvious bias toward publications in ISI journals, and an obvious bias away from writing books or publishing scholarly research in any form except for ISI journal articles.

This bias may result in some departments being ranked lower in publications than if books and other outlets were considered. When one recognizes that the NRC criteria are often factored into promotion, tenure, and hiring decisions, it creates a bias against scholars who spend their time writing books, and points researchers toward the production of journal articles rather than books. Certain types of ideas and analysis are better presented in books, where the author has room to present a more detailed argument. The ISI/NRC metric discourages those forms of scholarship.⁸ Certain fields may produce research more conducive to book-length treatment (economic history may be an example), and scholarship in those fields may be discouraged relative to other areas. At a minimum, the

⁷ All ISI publications and citations are counted equally, so as far as “measuring” quality by publication and citation counts, a publication or citation either counts or it does not. But economists have a long history of rating their journals by quality too. See Brauniger and Haucap (2003, 178) who note, “To test for any prejudices in favour of theoretical and quantitative journals, Hawkins, Ritter, and Walter (1973) included two ‘dummy’ journals in their survey of economists. Their nonexistent *Journal of Economic and Statistical Theory* ranked in the top third of all journals surveyed (24th out of 87) while the equally non-existent *Regional Studies and Economic Change* ranked in the bottom third (59th out of 87).” It may be that similar biases exist in rating entire departments, but this conjecture is beyond the scope of this essay.

⁸ For example, economists working in the Austrian tradition seem to be more inclined to publish their ideas in books rather than journal articles, perhaps creating a bias against such scholars. There are two journals aimed explicitly at publishing Austrian economics, *The Review of Austrian Economics* and *The Quarterly Journal of Austrian Economics*, but neither is in the ISI/SSCI database. For obvious reasons, this could push young scholars interested in the ideas of the Austrian school, but hoping to achieve tenure at a research university, toward doing more mainstream types of research.

publication of scholarly research will be biased toward journal articles in ISI journals.⁹

The NRC methodology only counts articles that were published in the most recent five years,¹⁰ so their ranking is an indicator of recent scholarship rather than lifetime achievement. An argument can be made for this, in that it places younger scholars on a more equal footing with senior scholars. However, one of the stated purposes of the NRC study is to help prospective graduate students to get an idea about the quality of doctoral programs, and if so, a longer time frame might provide a better indication of the scholarly achievement of a department's faculty.¹¹

Citations

The same database used to count publications is also used to count citations. Each article in the ISI database has a list of citations associated with it, and that list is used as the starting point for obtaining the NRC citation count. While every citation in an article is listed in the database, the NRC citation count truncates this list of citations in two substantial ways. First, the NRC only counts citations to articles that are in journals in the ISI database; second, the NRC only counts citations to articles that were published in the past five years.

Figure 1 shows the peculiar method of NRC. In the left column are three sources of scholarly publications, an article in the *Quarterly Journal of Economics* (*QJE*), which is in the ISI's Social Science Citation Index (SSCI), an article in *Constitutional Political Economy* (*CPE*), which is not in SSCI, and book called "Book X." Assume that each of these three publications cites the three references in the right column of the figure: an article from the

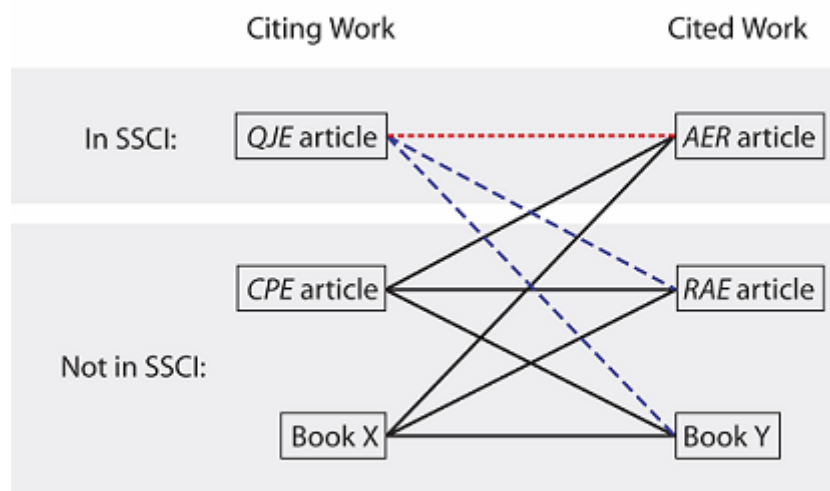
⁹ Klein and Chiang (2004b) suggest that there may be an ideological bias in the journals that are included in the ISI database as well.

¹⁰ The NRC 1995 report was actually completed in 1993, and used the five-year data period 1988, 1989, 1990, 1991, and 1992, as indicated on p. 3 of the NRC study. This was confirmed by Mr. Voytuk in a telephone conversation with Daniel Klein, 21 September 2004.

¹¹ The NRC presents publications per faculty member, rather than total publications, and who counts as a faculty member is determined by the institutions being ranked. They send a list of "program faculty" to the NRC; faculty on the list are counted, while those not listed are not. This clearly provides some discretion to the institutions, because it may not always be clear which faculty are a part of the doctoral program and which are not. I have no indication that universities have strategically excluded some faculty who are not publishing, but if the rankings are important to university administrators, the possibility is there.

American Economic Review (*AER*), which is in the SSCI, an article in the *Review of Austrian Economics* (*RAE*), which is not in the SSCI, and Book Y. Thus, three sources each make three citations, for a total of nine citations. Yet ISI/SSCI would count only three citations: those from the *QJE* article. NRC truncates even further by requiring that the cited work be in the ISI/SSCI, thus counting only one of those nine citations: the *QJE* citation of the *AER* article.

Citation counts do not count many citations



The diagram shows nine citations. SSCI counts only three (the dotted red and dashed blue citations). The NRC counts only one (the dotted red citation).

If citation count is intended to reflect the scholarly impact of a faculty member's work, there would seem to be no justification for not counting books, monographs, technical reports, and the like that are cited in ISI articles. The reason these citations are not counted has nothing to do with the merits of ISI journal articles relative to other cited publications, but is a function of the limitations of the ISI database. One problem with assigning citations to particular faculty members is that many people share names, so one cannot go by the author's name alone to determine who is being cited. The articles in the ISI database are cataloged by, among other things, the Zip Code of the author (NRC, 143), and the author's name

along with the Zip Code are used to assign citations under the NRC methodology. In order for the NRC to find the Zip Code of the cited author, the cited article must be in the ISI database. Thus, as with the publication count, if the work cited is a book, monograph, or article that is not in an ISI journal, it will not count as a citation because there is no Zip Code associated with the publication to identify the author. Part of the problem is that with so many faculty and publications, it is not feasible to count citations any way other than by using a computerized database, and one can utilize only the data that are cataloged. As the study notes, "The result of this matching was the identification of approximately 1 million publications that could be credited to the program faculty in the study." With so many publications, it would be infeasible to try to assign authorship to those works cited without a Zip Code in the ISI database.

Certainly in the case of some authors who have written frequently-cited books (e.g., Nobel laureates Douglass North and James Buchanan), their NRC citation counts seriously underrepresent the number of times they actually are cited relative to scholars whose cited works are published in ISI journals. Checking the ISI database to count all citations rather than just those to ISI articles is a relatively easy task, so for any one individual, one could correct for this oversight. However, if an institution's goal is to increase its faculty citations in the next NRC study, citations to articles in ISI journals will count more in that institution's incentive system.

Another problem with citation counts as an indicator of faculty quality is that some fields tend to cite more literature, and to cite more broadly, than others. For example, articles in macroeconomics tend to have fewer citations than those in labor economics, and articles in macroeconomics tend to cite a smaller set of key papers while articles in labor economics tend to cite a broader set of work to give a better feel for the empirical findings of past studies. The use of citation counts as an indicator of faculty quality tends to favor some fields of economics over others, so for hiring, promotion, and tenure, there may be a bias in economics departments favoring scholars in certain fields.

According to what the NRC study says, it appears that all citations appearing in ISI articles in the past five years to articles in ISI journals are counted, whenever the cited articles were published. However, the administrators at my university say that the NRC methodology only counts citations in ISI articles published in the last five years to articles published in ISI journals in the past five years, and this was confirmed by the NRC's Mr. James

Voytuk.¹² Thus, the NRC citation method is even narrower than the narrowness illustrated in Figure 1. If the cited article was published more than five years ago, it does not count as a citation in the NRC methodology. For purposes of annual evaluations, one thing we (economics faculty at Florida State University) report is a citation count done following the NRC methodology. The reason for not counting citations to articles published more than five years ago appears to be because the NRC used the same database of articles to count publications and citations, so articles older than five years would not have appeared in their database.

As with publications, this five-year truncation levels the playing field between junior and senior faculty. But if a department's prestige is based partly on faculty who have written classic articles, those articles would not be included in the NRC citation count, so the count would understate the prestige of such departments. The citation of older articles would suggest that those articles have stood the test of time, whereas newer articles may not, so if anything, it would appear better to bias the citation count in favor of older articles rather than to exclude them.

Considering the publication process in economics, it is difficult to get many articles cited in this time frame. After an article is written and submitted to a journal, the review time, the likelihood that the article will be rejected from a journal and submitted to a second one, and the frequency with which even accepted articles go through a process of revision, may take several years. Then there is often a publication lag of a year or more, so if a manuscript submitted for publication cites a newly published article, that article may be several years old before the manuscript citing it is published as an article. This lag would favor top departments in the citation count process, because it is more likely that an author in a top department would have colleagues that might cite their forthcoming work, allowing more of their work to be cited within that five-year window. By the time the citing article is published, the cited article may also be, and the citation can then be changed to list the final publication data.

Top departments may be favored for another reason. If scholars are evaluated by their citation counts, individuals can help each other in this regard by citing each others' work. This works best when those citing each others' work publish frequently in ISI journals, and because publication counts are higher in top-rated departments, those departments have an edge if this type of reciprocal citations occurs. This would not have to be an overt conspiracy to take advantage of the NRC methodology. People might

¹² In telephone conversation with Daniel Klein, 21 September 2004.

cite those they know well, such as their colleagues, and also might cite people in the hope that those authors will notice their work and reciprocate. That is more likely to happen when there is a personal connection. The Lucas (1976) critique applied to macroeconomic policy may also apply to citation counts. When citations are first counted, they may be a useful indicator of the influence of the author's work, but once people know they are used this way, the indicator loses its usefulness as authors exploit the system.

Finally, note that citations count the same regardless of the reason the work is cited. If an article has an error in it that is pointed out in a later work, or if the work is cited as an example of shoddy scholarship, it still counts as a citation.

Grants

Another metric in the NRC data is federal grants. The logic of using grants is that in order to get grants, researchers must submit their grant applications in a peer review process that evaluates the quality of the proposed work. The NRC does not use all grants, however, but only federal grants from the National Institutes of Health, the National Science Foundation, and the Department of Defense. As the study notes (NRC, 37), these agencies accounted for nearly three-quarters of all federal grants over the study period. The stated reason for limiting grant awards to these federal grants, however, is data availability. The principal investigators for these grants can be provided by the federal government and matched to the list of faculty supplied by the institutions. The NRC only counts federal grants because data are readily available from federal agencies (NRC, 37), and does not argue that federal grants are a better indicator of research quality than grants from private organizations or states, or that federal grants from these particular agencies are a better indicator than grants from other federal agencies.

The argument for including this measure in the NRC departmental rankings is that there is a positive correlation between federal grants to faculty and the quality rankings generated by the survey. Yet there are so many fundamental questions regarding this metric that one hardly knows where to start. The most obvious is that private grant support would seem to be at least as good an indicator of the quality of the recipient's research. To omit private grants obviously biases the rankings in favor of those who receive government versus private support. It also biases the incentives of

researchers toward seeking federal government support for their research rather than private support. This may affect the type of research one pursues. The National Institutes of Health and the Department of Defense have particular research interests that they want to fund, for example.

When one looks at the source of federal grants, this indicator of research quality can provide poor incentives. For example, the National Institutes of Health provides the bulk of their research dollars to universities with medical schools, so universities with medical schools will have an advantage in attaining federal grants. While grants are totaled by department, the synergy created between units can help generate grants. For example, it is easy to see that an economics department in a university with a medical school will be better positioned to obtain grants from the National Institutes of Health. This metric may bias the rankings of departments in favor of those that are at a university with a medical school, and may even create an incentive for a university to create a medical school in order to increase its federal funding, and thereby increase its NRC ranking.¹³

State and local government grants are not included in the NRC measure. This creates another bias, in that national issues are favored over local issues, and because of that, academic research in economics may tend to favor more centralized government activity over decentralized activity. Perhaps another factor is that the top universities, who have input into designing the methodology, tend to garner a larger percentage of federal grants, so federal grants is a measure that reinforces the perception that these departments have the top faculty.

Another problem with the idea that the amount of grant funding is a measure of the quality of a faculty is that it creates a bias toward undertaking more expensive research. Normally, in the private sector of the economy, if something can be accomplished more cheaply, that is viewed as desirable; yet, using grant funding as an indicator of quality, the more expensively one can undertake research, the better that will be for a department's ranking.¹⁴ Admittedly, there is some logic to using dollars as the metric from the standpoint of a university administration, because

¹³ At Florida State University, the administration is very conscious of the NRC rankings and encourages its faculty to get federal grants. Florida State University also established a medical school in 2000. While I have no evidence that the medical school was created as a method of increasing the university's federal grants, establishing a medical school may be a way to increase its performance in an area that the NRC finds is correlated with higher rankings of faculty quality.

¹⁴ Toutkoushian et al. (2003, 124) note that the level of grants may be a better measure of inputs into the production of research than the research output that is produced.

money is fungible, and grant money can be used for overhead, funding graduate students, and other purposes. From a social standpoint, however, less expensive research should be preferred to more expensive research, and it would seem that researchers should be praised for reducing the cost of their research. Instead, they are rewarded for increasing its cost.

Note that it is the dollars that are counted, not the number of grants. If the logic behind using grants as a measure of quality is that grants go through a review process that judges the quality of the research, number of grants would seem to be a better measure of quality than total grant dollars. From a quality standpoint, it would seem that five separately-reviewed grants for \$50,000 had more peer review, and should count for more than one \$250,000 grant-funded project.

The incentive structure creates some pressure for any researcher to seek out grant money, and especially federal grant money. It also may bias a department's interests toward those areas that are more expensive to undertake research, if that research might attract grant money. For example, experimental economics and computational economics are two areas in which it is relatively expensive to undertake research because of the cost of running experiments and the cost of computer equipment. These areas are frequently funded by National Science Foundation grants, making them good areas in which to specialize if a department wants to increase its federal grants. In this way, this criterion may bias the direction of research in economics in those departments that want to increase their standing in the NRC rankings.

Another issue related to federal grants is that the money to fund them comes from taxpayers who are forced to tender their tax dollars to the federal government. If my research depended on money that was forcibly taken from others, at the very least I would not be proud of this particular aspect of it, yet academic recipients of federal grants appear completely unapologetic of the fact that their work is funded by money taken from taxpayers against their will. Indeed, they even seem proud of it. At least with private funding, those who supply the grants do so voluntarily, and for this reason alone it would appear to me that foundation funding is far superior—on moral grounds alone—to government funding.

THE IMPACT OF THE NRC RANKINGS ON ECONOMIC RESEARCH

The NRC is working on an update of their rankings of doctoral programs, and because administrators at some universities want to see their doctoral programs rise in the rankings, the NRC rankings have an impact on economic research. I see this first-hand at my own institution, and colleagues in other mid-level economics programs tell me they see the same emphasis on the NRC rankings at their universities.¹⁵ Surely it is not unreasonable for university administrators to take account of these rankings, and to want their departments to improve in the rankings. The rankings are very visible to the academic community, and higher rankings provide more prestige to the university and to its administrators. Administrators cannot be intimately familiar with their doctoral programs that are well outside their fields of expertise, so it is natural for them to rely on this kind of outside ranking as at least a partial indicator of the quality of their programs. But by so doing, they alter the research incentives, hiring, promotion, and tenure decisions in their economics departments, even if that is not their intention. The NRC metrics are used as a measure of faculty performance simply because the information is collected and reported, and the incentives they imply affect the direction of research in economics. The methodology in the new updated study may differ from that in the current study. Yet because the new methodology is unknown at this time, the things that “count” right now are those used in the 1995 study.

In my own department, faculty are asked to supply a count of their publications, citations, and federal grants following the NRC methodology as a part of the annual evaluation process. This information is then passed on to the provost, who keeps track of faculty performance using the NRC methodology. These data are only a small part of the information on which faculty evaluations are based, and the provost has never said that he wants faculty to reorient their work so that they will add to the department’s NRC scores, but those NRC criteria take on an added importance because of the clear institutional goal to move up in the NRC rankings. They provide some incentive for faculty to orient their work so that they can get publications,

¹⁵ It is perhaps worth a footnote to remark that if quality is measured by these rankings, it is a zero-sum game among universities. If universities want their departments to move up in the NRC rankings, the successes of some automatically generate failures for others.

citations, and grants. Perhaps more significantly, they provide the incentive for the department to hire people who can do well in those areas.

The NRC ranks departments, not individual faculty.¹⁶ Individual faculty may be judged based on how they add to the NRC numbers, but a good way to “improve” a department, according to the NRC criteria, is to hire people who excel in those things. This may increase a department’s ranking (but remember, the ultimate rank order is a subjective evaluation, not these objective measures), and may also take some of the pressure off existing faculty to orient their work that way. The NRC rankings create an incentive to hire a particular kind of scholar. Doing interesting work, being intelligent, and working hard are all good things, but if they lead to publications, citations, and grants, as measured by the NRC, that is even better. As a result, departments that have their eyes on the NRC rankings will tend to become more homogeneous because they benefit from hiring in fields that tend to get more citations or grants. Departments may avoid individuals whose work is somewhat outside the mainstream because such work is less likely to be cited, or who may have good ideas but a slow publication rate, and the diversity of ideas in economics departments will likely decline as a result.

Thirty years ago, economics departments at places like UCLA, the University of Washington, and Virginia Tech had identities that set them apart from other departments. Now the research coming from those departments is much more similar to the research done in departments that rank above them. Among all the top-ranked departments, there is much more homogeneity today than there was decades ago.¹⁷ The homogenization of the discipline extends beyond schools of thought to the narrowing of fields of inquiry to focus on those that tend to get more citations, publications, and grants.

This emphasis on research and graduate education has also robbed academic economics of some of its relevance to real-world economic phenomena. Research departments give virtually no credit to faculty who write magazine articles, newspaper editorials, policy reports, or other material designed to enlighten the general public. Of course, one cannot blame the NRC departmental rankings for this, but the NRC rankings are a

¹⁶ One result of this is that there is no way for individual faculty members to look at their own numbers following the NRC methodology and identify any mistakes the NRC might have made in their calculations.

¹⁷ One department that stands out as having a unique identity is George Mason’s, which ranks 46th in the NRC study. See Landreth and Colander (2002, 6-8) for a discussion of the role of heterodoxy in the development of economic ideas.

part of a larger movement toward judging the quality of faculty in a very narrow way, such as by their academic journal publications, citations, and grants. The result is an academic discipline that is increasingly incapable of speaking to a general audience about real world problems.¹⁸

Looking at the big picture, the NRC ranking of economics departments is responsible for, at most, a small part of the changes in the nature of research done in economics departments, but it is part of a larger movement, as institutions want to improve their departments and try to do so by making their departments more like the top economics departments. This results in more homogeneous departments, and pushes individuals in those departments to orient their work more toward the mainstream so they can get publications, citations, and grants. As a result, the distinct identities that departments had 30 or 40 years ago are becoming more blurred. Without departmental rankings, departments might pride themselves in being among the best in a particular field, or for nurturing a unique approach or vision, but numerical rankings remove any value to uniqueness, and if anything, show that those unique characteristics do not pay off in terms of measured quality. The NRC rankings contribute to this homogenization of economics departments, and if one takes a Kuhnian (1962) approach to the advancement of the discipline, economics is likely to be worse off as a result.

REFERENCES

- Bauniger, Michael, and Justus Haucap.** 2003. Reputation and Relevance of Economics Journals. *Kyklos* 56(2): 175-198.
- Goldberger, Marvin L., Brendan A. Maher, and Pamela Ebert Flattau, eds.** 1995. (cited in the text as "NRC") *Research Doctorate Programs In the United States: Continuity and Change. Committee for the Study of Research-Doctorate Programs in the United States, National Research Council (NRC).* Washington, DC: National Academies Press.
- Jones, Lyle V., Gardner Lindzey, and Porter E. Coggeshall, eds.** 1982. *An Assessment of Research-Doctorate Programs in the United States.*

¹⁸ Years ago I had a colleague who, over lunch, told a story of a question asked by a student in his intermediate microeconomics class. The student asked what relevance the material they were covering had to the real world, and my colleague reported to us with delight that he told the student, "This is a theory class. If this material has any relevance to the real world, it is just a coincidence."

Washington, DC: National Academy Press.

Hawkins, Robert G., Lawrence S. Ritter, and Ingo Walter. 1973. What Economists Think of Their Journals. *Journal of Political Economy* 81(4): 1017-1032.

Klein, Daniel B., with Eric Chiang. 2004a. Citation Counts and SSCI in Personnel Decisions: A Survey of Economics Departments. *Econ Journal Watch* 1(1): 166-174.

———. 2004b. The Social Science Citation Index: A Black Box—with an Ideological Bias? *Econ Journal Watch* 1(1): 134-165.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Landreth, Harry, and David C. Colander. 2002. *History of Economic Thought* 4th ed. Boston: Houghton Mifflin.

Lucas, Robert E., Jr. 1976. Econometric Policy Evaluation: A Critique. In *The Phillips Curve and Labor Markets*, ed. Karl Brunner and Alan H. Meltzer. Amsterdam: North Holland, 19-46.

NRC. 1995. National Research Council: See Goldberger et al. above.

Toutkoushian, Robert K., Stephen R. Porter, Cherry Danielson, and Paula R. Hollis. 2003. Using Publications Counts to Measure an Institution's Research Productivity. *Research in Higher Education* 44(2): 121-148.

ABOUT THE AUTHOR



Randall G. Holcombe is DeVoe Moore Professor of Economics at Florida State University. He received his Ph.D. in economics at Virginia Tech, and taught at Texas A&M University and Auburn University before coming to Florida State in 1988. Dr. Holcombe is also Senior Fellow at the James Madison Institute, a Tallahassee-based think tank that specializes in issues facing state governments, and is a member of Governor Jeb Bush's Council of Economic Advisors. He is the author of ten books and more than 100 articles published in academic and professional journals. His email address is holcombe@garnet.acns.fsu.edu.

ECONOMICS IN PRACTICE

Economics without Entrepreneurship or Institutions: A Vocabulary Analysis of Graduate Textbooks

DAN JOHANSSON*

[Abstract, Keywords, JEL Codes](#)

TODAY, MANY VIEW ENTREPRENEURSHIP AND ECONOMIC freedom as the remedy for unemployment and low economic growth, not least in the European welfare states with permanently high unemployment and lagging economic growth. Economists have an important role to fulfill as policy advisers and civil servants. Do today's Ph.D. programs in economics give researchers adequate training in addressing questions concerning entrepreneurship?

Economics is a heterogeneous discipline with numerous traditions, each based on a cluster of theories. Each theory uses ideas, schemes, and assumptions. Different theories often give rise to opposing views on the importance of a problem, how the problems should be formulated, what methods should be applied, and what policy judgments to make.

*The Ratio Institute, Stockholm, Sweden.

I am grateful for comments on a Swedish version of this paper from Niclas Berggren, Gunnar Eliasson, Lennart Erixon, Magnus Henrekson, Rolf Henriksson, Per Hortlund, Richard Johnsson, Henrik Jordahl, Nils Karlson and participants in a workshop during the Small Business Days in Örebro arranged by the Swedish Foundation for Small Business Research. The article is based on information provided on the Internet by Swedish universities about the literature used in their Ph.D. programs. The information is easily accessible, but in case of ambiguities, the one responsible for the program, course, etc. has been contacted. I am grateful for the swift and complete replies of those contacted.

Theories are presented in textbooks. A textbook's index includes key words that indicate the structure of the theory, its method, and which problems it identifies as important. Words not appearing in a textbook's index are words not important in the theoretical structures within the book. We get an idea of what the textbook's theories do *not* consider to be important or have not yet captured. What is *not* written could be at least as telling as what is written.

My purpose is to investigate whether entrepreneurship-rich and institutions-rich theories are part of economics Ph.D. programs. I investigate whether key words appear in the index of the textbooks used in Sweden's economics Ph.D. programs.

The investigation covers all economics Ph.D. programs in Sweden for the academic year 2003-04. The investigation is not specific to Sweden, however, because Ph.D. programs in Sweden are a lot like programs in the United States. Nearly all of the textbooks examined are written by economists in the United States. The textbooks in Sweden are books familiar to Ph.D. students in the United States and elsewhere. So the Swedish aspect of this investigation is inessential. The investigation treats the dominant mainstream style of Ph.D. program, regardless of where on the globe it is situated.

The investigation covers the required courses in microeconomics and macroeconomics and courses in industrial organization (I/O). The required courses in microeconomics and macroeconomics present the theoretical foundation that everyone is supposed to know. Industrial organization is about industrial structure, competition, and development, so here especially students ought to encounter theories involving entrepreneurship and institutions.

Textbooks represent received theory, while articles are developments of theory and may or may not be fully received at a later point in time. Thus, articles are not included in the investigation.

THE DUAL LACUNAE: ENTREPRENEURSHIP AND INSTITUTIONS

The terms naturally break down into dual sets. One deals with knowledge and discovery: *entrepreneur*, *innovation*, *invention*, *tacit knowledge*, and

bounded rationality. The other deals with social rules: *institutions*, *property rights*, and *economic freedom*.

Entrepreneurship

In the history of economic thought, the entrepreneur and entrepreneurship have often been at the very center of analysis. Entrepreneurship was already discussed during Antiquity. The French 18th century economist Richard Cantillon was the first to integrate the entrepreneur into economic theory. Cantillon defined the entrepreneur as the one who took on business risk and took initiative to exploit business opportunities (Hebert and Link 1989). It was in connection with the growing dominance of the mathematical approach that the entrepreneur was removed from “mainstream economics”. This disappearance has been much noted. Just a few of the authors who explore the eradication the entrepreneur, usually indicting modern economics for it, include Schumpeter 1942: 86, Baumol 1968, Casson 1982, Barreto 1989, Hebert and Link 1982, Kirzner 1973: 26-26, Blaug 1986 (chap. 12), and Machovec 1995.

The entrepreneur plays a fundamental role in Austrian, Institutional and Schumpeterian theory, theories outside the mainstream paradigm. However, there is no universally accepted definition of the entrepreneur or of the entrepreneurial function. Seminal contributions have been made by Knight (1921), who defines the entrepreneur as the one who takes on genuine uncertainty,¹ and Kirzner (1973, 1997), who defines entrepreneurship as the faculty of discovering pure profit opportunities.

But, perhaps, Schumpeter has had the largest influence on today's research on the role of entrepreneurs. For Schumpeter, entrepreneurs generate and use new knowledge about how to better satisfy consumers in more efficient ways, driving economic development. He distinguishes between *invention* (coming up with a novel idea) and *innovation* (putting the invention to work). The entrepreneurial function is realized in innovation, actually introducing the invention into the economic system. This function is fundamental. Schumpeter lays out five broad categories of innovations:

¹ Risk is defined as a random event with a known distribution, while genuine uncertainty is a random event with an unknown distribution. The critical difference is that risk is insurable, while uncertainty is not.

1. The introduction of a new good - that is one with which consumers are not yet familiar - or of a new quality of a good.
2. The introduction of a new method of production, that is one not yet tested by experience in the branch of manufacture concerned, which need by no means be founded upon a discovery scientifically new, and can also exist in a new way of handling a commodity commercially.
3. The opening of a new market, that is a market into which the particular branch of manufacture of the country in question has not previously entered, whether or not this market has existed before.
4. The conquest of a new source of supply of raw materials or half-manufactured goods, again irrespective of whether this source already exists or whether it has first to be created.
5. The carrying out of the new organisation of any industry, like the creation of a monopoly position (for example through trustification) or the breaking up of a monopoly position. (Schumpeter 1934, 66)

Schumpeter stresses the importance of the organizational innovations bookkeeping and the stock company for the growth of the West. In fact, the more or less spontaneous development of private property rights in the West, which laid the foundation for its leading position (e.g. North and Thomas 1973; Rosenberg and Birdzell 1986), can be regarded as institutional innovations. It may be argued that institutional innovations are fundamental for technical ones. The actions of the entrepreneurs induce, in Schumpeter's (1942) words, creative destruction; old businesses are challenged by, and eventually replaced by, new ones.

According to this tradition, the entrepreneur can be seen not only as a factor of production, but as the most important factor of production. The reason is that he or she allocates all factors of production, including his own energy, attention, and vision, which makes it very special (Pelikan 1993). It works as a lever on the rate of return of all factors of production.

The entrepreneurial faculty is scarce and unequally distributed among the population, in quantity as well as in quality. Every entrepreneur is *boundedly rational*, i.e. he has a limited capacity to analyze and act on information (Simon 1955, 1990). Important parts of the faculty are *tacit*, impossible to articulate (Polanyi 1967). There are a limited number of

entrepreneurs who can carry out a limited number of entrepreneurial activities.

Institutions

It matters whether entrepreneurship is active, and if so, whether it is used productively, unproductively or destructively (Baumol 1990, Bhagwati 1982, Murphy *et al* 1991). The vitality of entrepreneurship relates directly to our second set of terms: *institutions*, *property rights*, and *economic freedom*.

Society's institutions – the rules of the game – largely determine the incentives of the entrepreneurs and thereby guide their actions. Private property rights are one of the most important institutions. The institutions to a large degree correspond to the degree of economic freedom, for instance, freedom of enterprise, the right for an individual to be an entrepreneur at all.

The clearest exponent of institutional theory and the importance of economic freedom is probably Adam Smith. The principal policy answer Smith gives to his query about the causes of the wealth of nations is economic freedom and the security of property rights. Boiled down to a single message, Smithian growth theory says freedom causes growth.²

There are many strands of institutional theory in the Smithian vein. Here I mention just a few. In the tradition of Ronald Coase, Armen Alchian, and Harold Demsetz, many property-rights economists like Terry Anderson and P. J. Hill interpret economic developments with the logic of property rights. Many economic historians like Robert Higgs and Douglass North make property rights and institutions the cornerstones of their historical explanations. Many policy economists like Sam Peltzman do serious empirical research on how regulations attenuate property rights and affect activity. Many Austrian, Public Choice, and New Institutional economists interpret economic topics with the logic of property rights and freedom of contract. These economists use words like *property rights* and *freedom*, not as policy judgments but as *analytic categories*.

Institutional theory looks upon growth as a process of knowledge accretion driven by entrepreneurs, whose behaviors are conditioned by institutions in general and by private property rights in particular (Kasper

² Smith identifies various factors that cause growth, and explains, in terms of other factors, *why* freedom causes growth, and even says that in some exceptional cases freedom should be contravened. But the main theory is that freedom causes growth.

and Streit 1998). Recent empirical support for the importance of institutions comes from the fall of the planned economies and other full-scale “experiments” where countries have applied different growth strategies, systematic analyses of the question why economic growth does not take off in some developing countries (e.g. De Soto 2000), and extensive economic-historical studies (e.g. North and Thomas 1973, Rosenberg and Birdzell 1986, Mokyr 1990).

The Duality between the Lacunae

It is no coincidence that dominant mainstream economics has the dual lacunae of entrepreneurship and institutions. Equilibrium thinking is essentially a system of mathematical functions. The mathematical representation of the theory rests on a number of axioms. Barreto (1989) writes:

The confrontation between the basic axioms and the entrepreneur leaves two possibilities: to accept the entrepreneur and reject the modern theory of the firm, or to reject the entrepreneur and maintain allegiance to the modern theory of the firm. . . . Simply put, entrepreneurship is above ‘formalization’—it cannot be neatly packaged within a mechanistic, deterministic model. Importantly, the choice is an ‘either-or’ proposition; there is no happy medium. The corner solution which economic theory has chosen is consistency and for this reason the entrepreneur disappeared from microeconomic theory. (Barreto 1989: 115, 141)

Analytically, all options are fully specified within a closed system, and the whole terminology of property rights is out of place. Entrepreneur-rich and institutional-rich traditions allow for actors to come up with creative action, interpretational breakthroughs. In this context, it is important to be able to speak of kinds of rules that constrain behavior (rules against stealing, for example) yet leave the door open for creative developments. Market entrepreneurship is transcendent action within a social framework of property rights. When economics cast its fate with equilibrium analysis, it made analysis of both entrepreneurship and institutions difficult.

UNIVERSITIES AND TEXTBOOKS

The investigation covers the 14 economic departments evaluated by Sweden's National Agency for Higher Education (*Högskoleverket* 2002): Göteborg University, Jönköping International Business School, Linköping University, Luleå University of Technology, Lund University, Stockholm University, Stockholm School of Economics, Swedish University of Agricultural Sciences (two departments, one in Uppsala and one in Umeå), Umeå University, University College of Dalarna, Uppsala University, Växjö University and Örebro University.

Several departments are too small to give the required courses in microeconomics and macroeconomics or courses in industrial organization. The students at the small departments may take them at larger departments.³ Stockholm School of Economics and Stockholm University have a joint program, Stockholm Doctoral Program in Economics, Econometrics and Finance (SDPE).

Hence, it is the large universities in Lund, Göteborg, Stockholm, Uppsala and Umeå that offer a complete course program. Jönköping also has a complete program. The requirements for a Ph.D. degree are similar for the different universities. The requirements encompass 160 credits ("points" in Swedish). Each credit is said to correspond to one week of full-time studies. With the exception of Lund, the credits are divided entirely between course work and writing the dissertation.⁴ Generally, the students begin by taking the required courses, thereafter the rest of the courses, and finally write the dissertation.⁵

Half of the courses, 40 credits, are required: 10 credits in microeconomics, macroeconomics, econometrics and mathematics, respectively, except in Lund where 30 credits are required (10 credits in microeconomics, macroeconomics and econometrics, respectively).⁶

³ The University College of Dalarna cooperates with Uppsala University, Linköping University cooperates with different universities, Luleå University of Technology cooperates with Umeå University, Swedish University of Agricultural Sciences cooperates with Umeå University and with Uppsala University, and Örebro University cooperates with Uppsala University.

⁴ In Lund, the courses comprise 70 credits and the writing of the dissertation 90 credits.

⁵ In Jönköping, the students start to write the dissertation at the same time they start to read the required courses.

⁶ In Lund, there are additional 15 credits in mathematical and statistical methods, which practically all Ph.D. students take. Together the 45 credits comprise the core courses. It is also possible for the students in Lund to exchange 5 credits in macroeconomics or

The studies are focused on mainstream economics. The theory is expressed in mathematical terms, the analysis is technical, and the students are trained in expressing the theory in mathematical form and to solve systems of equations. The empirical analysis focuses on econometrics and formal methods. It generally uses aggregated data and existing data sets. Surveys, case studies and interviews are uncommon. Passing exams and succeeding in a career depend on the student's ability to command and use the mathematical or econometric techniques. The requirements are high and competition is intense. Students can ill afford to pursue socially relevant applied research or to participate in or even follow current policy debate (Boschini *et al* 2004).

In total, 20 textbooks (different editions are counted as one book) are covered by the investigation, covering more than 11,000 pages. The texts are listed in Table 1. A few books dominate the education. Mas-Colell *et al* (1995) is the most commonly used textbook in microeconomics and is used in all courses in Micro I as well as in Micro II, Jönköping excepted. Varian (1992) is the second most used textbook in microeconomics. Romer (varied editions), Barro and Sala-i-Martin (varied editions) and Obstfeld and Rogoff (1996) dominate macroeconomics. Tirole (1989) is the main textbook in industrial organization. Although I have not done a study of Ph.D. programs in the United States, it is my strong impression that such programs have these same books as leading texts.

econometrics for courses more relevant for the dissertation. In Jönköping, the obligatory credits in mathematics and macroeconomics are reduced with two credits each to create room for a required course in the history of economic thought (4 credits).

Table 1 Universities, Courses, and Textbooks—academic year 2003-04

Ph.D. Program	Micro I	Micro II	I/O
Lund	Gibbons (1992) Mas-Colell et al. (1995)	Mas-Colell et al. (1995)	-
Göteborg	Jehle and Reny (1998) Mas-Colell et al. (1995) Varian (1992)	Jehle and Reny (1998) Mas-Colell et al. (1995) Varian (1992)	Tirole (1989)
Jönköping	Luenberger (1995)	Chambers (1988) Pollak and Wales (1996)	Buckley and Michie (1996) Schmalensee and Willig (1989)
Stockholm	Mas-Colell et al. (1995) Jehle and Reny (2001)	Mas-Colell et al. (1995) Laffont and Martimort (2002)	Tirole (1989)
Uppsala	Mas-Colell et al. (1995) Varian (1992)	Mas-Colell et al. (1995) Varian (1992)	-
Umeå	Varian (1992)	Mas-Colell et al. (1995)	-
Ph.D. Program	Macro I	Macro II	
Lund	Romer (1996)	Romer (1996) Obstfeld and Rogoff (1996)	
Göteborg	Barro and Sala-i-Martin (1995) Romer (2001)	Barro and Sala-i-Martin (2004) Romer (2001)	
Jönköping	Obstfeld and Rogoff (1996)	Obstfeld and Rogoff (1996)	
Stockholm	Barro and Sala-i-Martin (1995) Blanchard and Fischer (1989) Ljungqvist and Sargent (2000) Obstfeld and Rogoff (1996) Romer (1996) Sargent (1987) Stokey and Lucas (1989) Walsh (1998)	Barro and Sala-i-Martin (1995) Pissarides (2000) Ljungqvist and Sargent (2000)	
Uppsala	Barro and Sala-i-Martin (1999), Romer (2001)	Romer (2001)	
Umeå	Barro and Sala-i-Martin (1995), Blanchard and Fisher (1989)	Blanchard and Fisher (1989)	

Note: Textbooks recommended as useful complementary literature are included. The exception is the course in industrial organization in the SDPE, in which several undergraduate textbooks were suggested. Industrial organization in Göteborg University was moved to the Fall 2004. The textbook listed here corresponds to last time the course was offered. Macro II in Uppsala University refers to single chapters in other textbooks as literature; those textbooks are not included. Macro I in Umeå University was postponed one

year because of a shift in the course program. In this case, the textbook listed here corresponds to the academic year 2002-03.

PRESENCE AND MEANING: A VOCABULARY ANALYSIS

A reference to an expression is counted in the following manner: “Innovation 64,” one reference, “Innovation 64-67,” one reference, “Innovation, 37, 64-67,” two references, etc.⁷

Table 2 Textbooks, concepts, presence, the academic year 2003-04

Textbook	Total Pp.	Total # refs.	Institutions/freedom ideas		
			Institution	Property rights	Economic Freedom
Barro and Sala-i-Martin (2004)	654	9	0	0	0
Blanchard and Fisher (1989)	650	0	0	0	0
Chambers (1988)	331	0	0	0	0
Gibbons (1992)	267	0	0	0	0
Jehle and Reny (2001)	543	0	0	0	0
Laffont and Martimort (2002)	421	2	0	1	0

⁷ I include all variants, like *entrepreneur*, *entrepreneurial* and *entrepreneurship*. Several textbooks refer to *Technological innovation*. I have counted *technological innovation* as *innovation*. I have also included references to *process innovation* and *product innovation* under *innovation* (this applies to Tirole 1989). I have not included *knowledge*, *knowledge accumulation*, *research and development*, *technical change*, *technological change* or *technology*. I have also excluded terms that may be regarded as synonymous, for instance *innovators* (one reference in Romer 2001), which could be interpreted as a synonym for entrepreneur, because of the indeterminacies involved in looking for synonyms.

Table 2 (continued)

Textbook	Total Pp.	Total # refs.	Institutions/freedom ideas		
			Institution	Property rights	Economic Freedom
Ljungqvist and Sargent (2000)	701	1	0	0	0
Luenberger (1995)	486	1	0	1	0
Mas-Colell <i>et al</i> (1995)	981	1	0	0	0
Obstfeld and Rogoff (1996)	804	0	0	0	0
Pissarides (2000)	252	1	0	1	0
Pollak and Wales (1996)	217	0	0	0	0
Romer (2001)	651	5	0	4	0
Sargent (1987)	510	0	0	0	0
Schmalensee and Willig (1989)	1555	64	1	0	0
Stokey and Lucas (1989)	588	0	0	0	0
Tirole (1989)	479	16	0	0	0
Walsh (1998)	528	4	4	0	0
Varian (1992)	548	1	0	1	0
Total:	11,166	105	5	8	0

Textbook	Knowledge/discovery ideas				
	Entrepreneur	Innovation	Invention	Bounded rationality	Tacit knowledge
Barro and Sala-i-Martin (2004)	0	9	0	0	0
Blanchard and Fisher (1989)	0	0	0	0	0
Chambers (1988)	0	0	0	0	0
Gibbons (1992)	0	0	0	0	0
Jehle and Reny (2001)	0	0	0	0	0

Table 2 (continued)

Textbook	Knowledge/discovery ideas				
	Entrepreneur	Innovation	Invention	Bounded rationality	Tacit knowledge
Laffont and Martimort (2002)	0	0	0	1	0
Ljungqvist and Sargent (2000)	0	1	0	0	0
Luenberger (1995)	0	0	0	0	0
Mas-Colell <i>et al</i> (1995)	1	0	0	0	0
Obstfeld and Rogoff (1996)	0	0	0	0	0
Pissarides (2000)	0	0	0	0	0
Pollak and Wales (1996)	0	0	0	0	0
Romer (2001)	1	0	0	0	0
Sargent (1987)	0	0	0	0	0
Schmalensee and Willig (1989)	0	60	0	3	0
Stokey and Lucas (1989)	0	0	0	0	0
Tirole (1989)	0	16	0	0	0
Walsh (1998)	0	0	0	0	0
Varian (1992)	0	0	0	0	0
Total:	2	86	0	4	0

Note: Buckley and Michie (1996) is a compilation of articles that does not include any index. It is therefore excluded from the analysis. Schmalensee and Willig (1996) is not a textbook but a compilation of papers that illuminates particular themes. It is unclear whether it should be included or not, but I chose to include it because it contains an index. Several editions of Barro and Sala-i-Martin (1995, 1999, 2004), Jehle and Reny (1998, 2001) and Romer (1996, 2001) are used in the courses. The table only shows the latest editions. This does not affect the results.

Of the 19 leading textbooks, 16 contain five or fewer references to any of the entire set of eight terms. Eight of the leading textbooks contain no reference to any of them. Among the 19 books, only 2 references are made to *entrepreneur*-, only 5 to *institutions*, only 8 to property rights, and not

a single reference to *economic freedom*, *invention*,⁸ or *tacit knowledge*. It is quite obvious that economists have eradicated entrepreneurship and institutions from core Ph.D. training.

Moreover, in the textbooks where references are made, the references are usually few,⁹ and the meaning and significance of the ideas are lost, diluted, or distorted, compared to the entrepreneurship-rich and institution-rich theories. The reference to *entrepreneur* in Mas-Colell *et al* (1995)—one of two references made to the concept entrepreneur—is telling.

13.C.6 Consider a market for loans to finance investment projects. All investment projects require an outlay of 1 dollar. There are two types of projects: good and bad. A good project has a probability of p_G of yielding profits of $\Pi > 0$ and a probability $(1 - p_G)$ of yielding profits of zero. For a bad project, the relative probabilities are p_B and $(1 - p_B)$ respectively, where $p_G > p_B$. The fraction of projects that are good is $\lambda \subseteq (0, 1)$. ¶ Entrepreneurs go to banks to borrow the cash to make the initial outlay (assume for now that they borrow the entire amount). A loan contract specifies an amount R that is supposed to be repaid to the bank. Entrepreneurs know the type of project they have, but the banks do not. In the event that a project yields profits of zero, the entrepreneur defaults on her loan contract, and the bank receives nothing. Banks are competitive and risk neutral. The risk-free rate of interest (the rate the banks pay to borrow funds) is r . Assume that . . . (Mas-Colell *et al* 1995, 475)

The entrepreneur is not mentioned at all in the fundamental function she undertakes in Schumpeterian or Kirznerian theory, but could be any borrower at all. The same is true for the other reference, in Romer (2001: 394-398). The reference is made to *Entrepreneur-investor contracts*, i.e. a loan contract between a borrower (the entrepreneur) and the lender (the investor). In this case, it could also apply to any borrower, and the

⁸ Barro and Sala-i-Martin (2004) make one reference to *invention* under the topic of innovation. They use the terms synonymously, so I classify it as *innovation*.

⁹ The exception is Schmalensee and Willig (1989) with 60 references to innovation. On the other hand, this textbook makes no references to *entrepreneur* or *invention*.

entrepreneurial function is absent. The reference is made in a chapter entitled *Financial-Markets Imperfections*, since the actors in the model are assumed not to possess perfect information. This stands in sharp contrast to the traditions that have a developed theory about the entrepreneurial function. These theories, probably most accentuated in the Austrian tradition, attribute the success of the market economy to its ability to more effectively than competing economic systems generate and use new knowledge. It is in the nature of things that it is impossible for new insight to be available for all people at the same time. It is created in individual persons' brains. The entrepreneurial function is to identify and introduce new knowledge into the market, which is disseminated by imitation, experience, observation, and conversation. The adjustments of the economic actors to the new knowledge lead to creative destruction and economic transformation. It is the profits that in a first phase go to the entrepreneur that are the driving force behind economic development and economic growth. Competition and the process of the market economy may, according to this tradition, be compared to a procedure for the discovery and use of new knowledge (Hayek 1937, 1945, 1978).

Bounded rationality is referred to in two books, Laffont and Martimort (2002: 393) and Schmalensee and Willig (1989: 109-110, 138-139, 170-171). Only the latter uses the concept in its original sense.

Innovation is referred to in four textbooks and then tantamount to technical innovations, resulting from research and development.¹⁰ Organizational and institutional innovations are absent from the analysis. In traditions focusing on the entrepreneurial function, innovations are not driven narrowly by research and development, but by entrepreneurs pursuing new business opportunities. Furthermore, the textbooks do not distinguish between invention and innovation. According to Schumpeterian theory, research and development gives rise to inventions, not innovations. It is only when an entrepreneur commercializes the invention that it is proper to talk about innovations. The distinction has important implications for economic policy. Increased funding to research and development need not increase employment and economic growth if commercialization mechanisms function poorly, for example because the environment is unfavorable to entrepreneurship. In that case, an increased return on research and development can be achieved by improving the

¹⁰ Several of the textbooks make a reference to *Research and development* and/or *Technological innovation* under the concept *innovation*. Ljungkvist's and Sargent's (2001) reference to *innovation* is made to in time series representation.

environment of entrepreneurship, e.g. by decreasing taxes and removing restrictions. This is one explanation of the poor return on investments in research and development in Sweden (Henrekson and Rosenberg 2001).

Only five textbooks refer to *property rights*, and in those cases just in passing. This is a bit surprising considering the last few decades' research, especially in economic history and in institutional economics, showing that Smith was right about the establishment and protection of private property rights as a central factor in economic growth (e.g. Acemoglu *et al* 2001, 2004, Hall and Jones 1999, Mokyr 1990, North and Thomas 1973, Rosenberg and Birdzell 1986, Gwartney and Lawson 2004). Moreover, it is not the fundamental role of property rights for the working of the economy that are examined. Instead property rights are discussed in connection with market failures and external effects, i.e., not the general importance of property rights, but only in the case of environmental pollution, etc. (Luenberger 1995: 321, Romer 2001: 36-39 and Varian 1992: 435). Romer (2001: 116, 120, 121) also discusses property rights in connection with the creation of knowledge. Pissarides (2000: 194) mentions property rights in connection with job matching. Laffont's and Martimort's (2001: 373) reference is only a reference to others.

The results for *institution* are similar. Schmalensee and Willig (1989: 63-64) mention that institutions strive to reduce transaction costs. Walsh (1998: 160-162, 375-381, 380-381, 371-375) discusses the institutional set up for governing central banks. No textbook discusses institutions in terms of the "rules of the game" of society that govern the actions of economic actors and hence affects overall economic performance. The most commonly used textbooks in microeconomics (Mas-Colell *et al* 1995 and Varian 1992) and macroeconomics (Barro and Sala-i-Martin 2004, Obstfeld and Rogoff) do not refer to, for instance, North and Thomas (1973), Rosenberg and Birdzell (1986) or Mokyr (1990) who all, in comprehensive and well-known economic-historical studies, show that the economic success and political dominance of the Western World is due to the establishment and protection of private property rights and other market-conforming institutions.^{11, 12} A book like Barro and Sala-i-Martin (2004), entitled *Economic Growth*, hence, does not make any reference to entrepreneurship, institutions, private property rights, or to the economic-historical studies that have documented the importance of private property

¹¹ North received The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel in 1993.

¹² Romer (2001) refers to North (1981).

rights and other institutions for long-run economic growth. Tirole (1989), probably the most used textbook in industrial organization in the World, contains no references to *entrepreneur*, *institution* or *property rights*.

In the 19 books there are many references to Nash equilibrium, Bertrand equilibrium, Cournot equilibrium, Concave utility functions, Euler equations, etc.; concepts that on the other hand are not used at all in, e.g., Austrian, Institutional or Schumpeterian traditions. In these traditions, the concept of equilibrium is hardly seen as relevant or useful. Instead, they emphasize that actors have *disjoint* knowledge (that is, not merely asymmetric information, but asymmetric interpretations), the economy is a dynamic open-ended process in continuous change, and the scope and motivation for discovery is conditioned by the social rules.

SWEDISH REFLECTIONS ON THE PROBLEM

There are researchers who worry that today's Ph.D. programs in economics educate researchers unable to identify and analyze economically interesting problems with great relevance for society. Assar Lindbeck (2001) wants "two-leg" economists, i.e. economists commanding the formal mathematical analysis as well as being able to formulate and investigate important issues. Lars Calmfors (1996, 239-240) is of the opinion that the focus on mathematics and statistics makes new researchers untrained in solving real-world problems. He sees it as a risk that students interested in real-world problems are not going to succeed in the academic competition, being beaten out by those talented in technical crafts rather than relevant, meaningful knowledge and analysis. Lennart Erixon (2001, 317) regards this as a problem, not just for the economics profession, but more importantly for democracy: "The lack of new generalists is not just an internal problem but also a threat to democracy, if it prevents an independent professional elucidation of political decisions with great importance for the citizens' wealth."

Still, many researchers, e.g. both Calmfors and Lindbeck, think that mathematics and the training in formal methods provide a good, even necessary, basis for the Ph.D. students. Other researchers think otherwise and are of the opinion that it is the mathematical model-building dominating today's research that causes the problem. Professor emeritus Erik Dahmén is of the opinion that mathematically oriented economists are

“prisoners of the tool shed”. Their theory and method make them incapable of defining and analyzing economically relevant problems; they choose problems according to what their mathematical “tool-box” can handle and not according to what is important for society. Gunnar Eliasson writes:

From the classical [equilibrium] model only firms that do not belong to this world can be derived. This makes it useless as a theoretical foundation for studying management and information problems of firms. . . . I would propose to get the classical model removed from organization theory, from the theory of the firm, and as a foundation of macroeconomics. (Eliasson 1996: 23, 37)

This investigation implies that the theory underlying all Ph.D. programs in economics in Sweden excludes what chiefly explain economic growth and general wealth—entrepreneurship and private property rights. Then it is not a surprise, but rather natural, that the younger generation of economists do not participate more in the public debate. Their education is founded on theories and methods often useless in analyzing real-world issues.

Researchers who study technological development stress that development is carried out within, and are limited by, the design space of individuals from where they “get” ideas (Stankiewicz 2000). Similarly, if concepts like entrepreneurship and property rights are missing in the design space of economists, then those concepts will be excluded from the analysis. But is it possible for researchers to describe and analyze, for instance, the progress of the furniture industry or the progress of the computer industry, in a credible way, without taking account of the entrepreneurs Ingvar Kamprad or Bill Gates and the entrepreneurial function they have carried out, manifested in the founding and expansion of IKEA and Microsoft? Is it possible to grasp the development of Sweden’s industry excluding the inventors and entrepreneurs who once founded and developed the big firms that today comprise the country’s economic backbone? And in understanding the prominent Schumpeterian stories like Kamprad’s IKEA, we come to understand entrepreneurship in the general process of economic betterment, including all the small Kirznerian stories. We learn something that cannot be learned from a system of equations. Is it possible to analyze total employment and economic growth, the aggregated outcomes of the actions of individual persons and firms, without a theory of entrepreneurship? In what way does

omitting the entrepreneurial function from the analysis influence our understanding of enterprise, economic development and economic growth? Schumpeter (1942: 86) famously compared leaving the entrepreneur out of economic theory to leaving the Prince of Denmark out of *Hamlet*.

CONCLUSION

James Buchanan pointed out the same lacunae in equilibrium economics. In the Postscript of *What Should Economists Do?*, he included these two items in a list of points of what economics was failing to get right.

Economics involves actors. Without actors, there is no play. This truism has been overlooked by modern economists whose universe is people with passive responders to stimuli. If all are price-takers, who sets price? If all behavior is rationally responsive, how can change occur? How can entrepreneurship be modeled? Increasingly, I have come to the view that the role of entrepreneurship has been the most neglected area of economic inquiry, with significant normative implications for the general understanding of how the whole economy works. (Buchanan 1979, 281)

Economics is about a game within rules. Choices are made by actors, by traders, constrained within specifically determined 'laws and institutions,' a central emphasis of Adam Smith and one that has been lost to modern minds. (p. 281-82; italics in the original)

Buchanan's assessment of an economics lacking these insights is rather bleak.

I see a continuing erosion of the intellectual (and social) capital that was accumulated by 'political economy' in its finest hours. I look at young colleagues trained to master regression routines who are totally uninterested in, and incompetent to examine, elementary economic

propositions. . . . I see them compelled to utilize their considerable mental potentials resolving the escapist puzzles of modern mathematics. (pp. 279-280)

Inspection of the leading textbooks confirms Buchanan's remarks.

Speaking of industry in Sweden, Erik Dahmén says that the problem is not the industry we do have, but the industry we do *not* have. Similarly, the problem with economics education is not the training we do have, but the training we do *not* have. My conclusion, therefore, is that there is a need for economics Ph.D. training based on theories that incorporate entrepreneurship and institutions.

REFERCENES

- Acemoglu, D, Johnson, S and Robinson J.** 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review* 91(5): 1369-1401.
- Acemoglu, D, Johnson, S and Robinson J.** 2004. Institutions as the Fundamental Cause of Long-Run Growth. Working paper no 10481, NBER.
- Aghion P and Howitt P.** 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Barreto, H.** 1989. *The Entrepreneur in Microeconomic Theory: Disappearance and Explanation*. London: Routledge.
- Barro, R and Sala-i-Martin X.** 1995. *Economic Growth*. New York: McGraw-Hill.
- Barro, R and Sala-i-Martin X.** 1999. *Economic Growth*. Cambridge, MA: MIT Press.
- Barro, R and Sala-i-Martin X.** 2004. *Economic Growth*, Cambridge, MA: MIT Press.
- Baumol, W.** 1968. Entrepreneurship in Economic Theory *American Economic Review* 58(2): 64-71.

- Baumol, W.** 1990. Entrepreneurship: Productive, Unproductive and Destructive. *Journal of Political Economy* 98(5): 893-921.
- Bhagwati, J.** 1982. Directly, Unproductive, Profit-Seeking DUP Activities. *Journal of Political Economy* 90(5): 988-1002.
- Blanchard, J-O and Fisher S.** 1989. *Lectures on Macroeconomics*. Cambridge, MA: MIT Press.
- Blaug, Mark.** 1986. *Economic History and the History of Economics*. Brighton, Sussex: Wheatsheaf Books.
- Boschini, A, Lindquist, M, Pettersson, J and Roine J.** 2004. Learning to Lose a Leg: Casualties of PhD Economics Training in Stockholm. *Econ Journal Watch* 1(2): 369-379.
- Buckley, P and Michie J (ed.).** 1996. *Firms, Organizations and Contracts: A Reader in Industrial Organization*. Oxford, UK: Oxford University Press.
- Buchanan, J.** 1979. *What Should Economists Do?* Indianapolis, IN: Liberty Press.
- Casson, Mark.** 1982. *The Entrepreneur: An Economic Theory*. Oxford: Martin Robertson.
- Calmfors, L.** 1996. "Nationalekonomernas roll under det senaste decenniet – vilka är lärdomarna?", in Jonung, L (ed.), *Ekonomerna i debatten – gör de någon nytta?*. Ekerlids förlag, Stockholm.
- Chambers, R.** 1988. *Applied Production Analysis*. Cambridge, MA: Cambridge University Press.
- De Soto, H.** 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. London: Bantam.
- Eliasson, G.** 1996. *Firm Objectives, Controls and Organization. The Use of Information and the Transfer of Knowledge within the Firm*. Dordrecht: Kluwer Academic Publishers.
- Erixon, L.** 2001. "Finns det några debattörer för Ekonomisk Debatt?". *Ekonomisk Debatt* 29(5): 315-317.
- Gibbons, R.** 1992. *A Primer in Game Theory*. London: Harvester Wheatsheaf.
- Gwartney, James and Robert Lawson.** 2004. Economic Freedom of the World: 2004 Annual Report. Vancouver, BC: The Fraser Institute.

- Hall, R and Jones C.** 1999. Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics* 114(1): 83-116.
- Hayek, F.** 1937. Economics and Knowledge. *Economica* 4(1): 33-54.
- Hayek, F.** 1945. The Use of Knowledge in Society. *The American Economic Review* 35(4): 21-30.
- Hayek, F.** 1978. Competition as a Discovery Procedure. in Nishiyami, C and K Leube *The Essence of Hayek*. Stanford, CA: Hoover Institution Press.
- Hébert, R and Link, A.** 1982. *The Entrepreneur: Mainstream Views and Radical Critics*. Praeger Publishers.
- Hébert, R and Link, A.** 1989. In Search of the Meaning of Entrepreneurship. *Small Business Economics* 1(1): 39-49.
- Henrekson, M and Rosenberg, N.** 2001. Designing Efficient Institutions for Science-Based Entrepreneurship: Lessons from the US and Sweden. *Journal of Economic Perspectives* 26(2): 207-231.
- Högskoleverket 2002.** *Utvärdering av utbildning i nationalekonomi vid svenska universitet och högskolor*, Högskoleverkets rapportserie 2002:9 R.
- Jehle, G and Reny, P.** 1998. *Advanced Microeconomic Theory*. Reading, PA: Addison-Wesley.
- Jehle, G and Reny, P.** 2001. *Advanced Microeconomic Theory*. Boston: Addison-Wesley.
- Kasper, W and Streit M.** 1998. *Institutional Economics*. Cheltenham, UK: Edward Elgar.
- Kirzner, I.** 1973. *Competition and Entrepreneurship*. Chicago: The University of Chicago Press.
- Kirzner, I.** 1985. *Discovery and the Capitalist Process*. Chicago: The University of Chicago Press.
- Kirzner, I.** 1997. Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature* 35(1): 60-85.
- Knight, F.** 1921. *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin.

- Laffont, J-J and Martimort, D.** 2002. *The Theory of Incentives*. Princeton, NJ: Princeton University Press.
- Lindbeck, A.** 2001. Economics in Europe. *CESifo Forum* 2(1): 31-32.
- Ljungqvist, L and Sargent, T.** 2000. *Recursive Macroeconomic Theory*. Cambridge, MA: The MIT Press.
- Luenberger, D.** 1995. *Microeconomic Theory*. New York: McGraw-Hill.
- Machovec, Frank M.** 1995. *Perfect Competition and the Transformation of Economics*. London: Routledge.
- Mas-Colell, A, Whinston, M and Green, J.** 1995. *Microeconomic Theory*. Oxford, UK: Oxford University Press.
- Mokyr, J.** 1990. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford: Oxford University Press.
- Murphy, K, Shleifer, A and Vishny, R.** 1991. The Allocation of Talent: Implications for Growth. *The Quarterly Journal of Economics* 106(2): 503-530.
- North, D.** 1981. *Structure and Change in Economic History*. New York: W.W. Norton.
- North, D and Thomas, R.** 1973. *The Rise of the Western World: A New Economic History*. Cambridge, MA: Cambridge University Press.
- Obstfeld, M and Rogoff, K.** 1996. *Foundations of International Macroeconomics*. Cambridge, MA: The MIT Press.
- Pelikan, P.** 1993. Ownership of Firms and Efficiency: The Competence Argument. *Constitutional Political Economy* 4(3): 349-392.
- Polanyi, P.** (1967). *The Tacit Dimension*. London: Routledge & Kegan Paul.
- Pissarides, C.** 2000. *Equilibrium Unemployment Theory*. Cambridge, MA: MIT Press.
- Pollak, R and Wales, T.** 1996. *Demand System Specification and Estimation*. Oxford: Oxford University Press.
- Romer, D.** 1996. *Advanced Macroeconomics*. New York: McGraw-Hill.
- Romer, D.** 2001. *Advanced Macroeconomics*. New York: McGraw-Hill.

- Rosenberg, N and Birdzell, L.** 1986. *How the West Grew Rich: The Economic Transformation of the Industrial World*. New York: Basic Books Inc.
- Sargent, T.** 1987. *Macroeconomic Theory: Economic Theory, Econometrics and Mathematical Economics*. Boston: Academic Press.
- Schmalensee, R and Willig R (ed.).** 1989. *Handbook of Industrial Organization: Volume 1 and Volume II*. Amsterdam: North Holland.
- Schumpeter, J.** 1934. *The Theory of Economic Development*. New Brunswick: Transaction Publishers.
- Schumpeter, J.** 1942. *Capitalism, Socialism and Democracy*. New York: Harper & Row.
- Simon, H.** 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69(1): 99-118.
- Simon, H.** 1990. Invariants of Human Behavior. *Annual Review of Psychology* 4(1): 1-19.
- Stankiewicz, R.** 2000. The Concept of 'Design Space'. in Ziman, J (ed.), *Technological Innovation as an Evolutionary Process*. Cambridge, UK: Cambridge University Press, 234-247.
- Stokey, N and Lucas, R.** 1989. *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.
- Tirole, J.** 1989. *The Theory of Industrial Organisation*. Cambridge, MA: MIT Press.
- Tollison, R.** 1997. "Rent Seeking", i Mueller, D (red), *Perspectives on Public Choice: A Handbook*, Cambridge University Press, Cambridge, UK, s 506-525.
- Walsh, C.** 1998. *Monetary Theory and Policy*. Cambridge, MA: MIT Press.
- Varian, H.** 1992. *Microeconomic Analysis*. New York: W&W Norton & Co.

ABOUT THE AUTHOR



Dan Johansson received his Ph.D. from the Royal Institute of Technology in 2001. Since 2002 he has been a research fellow at the Ratio Institute in Stockholm. His research mainly deals with the role of entrepreneurs, the importance of new and small firms and the effects of institutions on entrepreneurship, firm formation and firm growth. His research, which is both applied and theoretical, currently focuses on the Swedish Growth School, a unique Swedish tradition influenced by among other things Austrian and early Schumpeterian theory. He has published in international and Swedish journals, co-authored and edited several books, and has taught at the Stockholm School of Economics and at the Royal Institute of Technology.

CORRESPONDENCE

Editors,

Deirdre McCloskey has complained about economists' mistaking statistical significance for significance. I thought the horse she was beating was dead already. But now she and Stephen Ziliak (in ["Size Matters" EJW August 2004](#)) demonstrate the beast is still breathing. For some reason McCloskey's logic hasn't penetrated.

Sometimes one good example succeeds where logic fails. It doesn't have to be a "real" example; a hypothetical should serve.

So I propose that every year, from a large population like that of the USA, a million forty-year-olds are selected at random and assigned, randomly, half to an experimental cohort and half to a control cohort, the former prescribed (and enforced) a 325 milligram aspirin tablet per day, the other half a placebo, both pills coated to avoid differences in taste or texture. As they die their age at death is recorded, and whether their death was related to coronary artery disease. Eventually we have recorded a few tens of millions of deaths and can identify any discrepancy in age of death, or cause of death, between the experimentals and the controls. I propose that the data support a positive benefit from the daily aspirin at a .0000001 level of confidence. If that's not convincing, keep going for another decade or two and the level goes to .0000000001. (Let me ignore how we enforce the regimen, and whether over half a century changes in diet, lifestyle, environment, etc., might cloud the results.)

What do we know about the benefits of a daily aspirin? McCloskey and Ziliak would say, "almost nothing." I'd say, only slightly different, "that there may possibly be some value in taking an aspirin per day, worth looking into." When we do look into it we may find that the daily aspirin's value, thought positive, is worth something less than the (small) cost of the daily aspirin, or even less than the trouble of pouring a glass of water and

remembering the daily aspirin. Of course it may make a whale of a difference. The McCloskey point is that the .00000001 or the .0000000001 doesn't tell us whether to take an aspirin.

Of course, there will be those among us who decide that aspirin is cheap and any positive result is sufficient, irrespective of magnitude. But if that's one's attitude, there is yet no demonstration that an aspirin a day makes a significant difference. We want the experimenters to look at their results and resolve the issue: Does aspirin make enough difference to take seriously? The .0000000001 doesn't tell me the answer.

Thomas C. Schelling
University of Maryland

ZILIAK AND MCCLOSKEY REPLY:

Editors,

We were pleased to read the letter by Tom Schelling, showing decisively with a new example that statistical significance is neither necessary nor sufficient for proving *economic* significance. Tom notes that the beast—The Standard Error—is still breathing. We can add a little precision: now it dominates in 80% or more of the published papers. We've got to kill it if economics is to progress.

We asked our old friend William Kruskal, a distinguished statistician and past president of the American Statistical Association, "How could the confusion of statistical significance for economic significance proliferate?" He replied, "Well, I guess it's a cheap way to get marketable results." Indeed. An economist couldn't put it better. Orley Ashenfelter told us recently that he tended to agree that the procedure is nonsense, but that young people have to run their careers. It's a wonderment that economists, even Orley, who teaches the stuff, do not yet feel ashamed.

Results of Tom's own study of empirical methods in the 1998 "Committee on Journals" suggest dissatisfaction with the publication standards of the *American Economic Review*. A third of the respondents to the Committee's survey said of the *Review* "there is too little empirical data." "Hardly any [said] too much . . . three-fifths [of the respondents said there is] too little

policy focus [too little emphasis on effect size and real-world importance], hardly any [said] too much.”

Significance testing is of course a big part of an empirical publication in the *AER*. Over 95% of the empirical papers published in it rely on significance testing to *some* degree; to rely on it to any degree is a plain error in elementary statistics. Schelling’s Committee did not look into significance testing. But given the response of the profession to the Committee’s findings (namely, none), and to our 1996 paper in the *JEL* at about the same time (next to none), we suspect nothing would have changed had a question on significance been added. Tom urges us to keep trying.

Raising the price of t is the ticket. Fisher’s 1925 publication of Student’s table of t was slowed down by a weak-armed assistant of Gosset, a young man who could only barely turn the crank of the calculating machine. Gosset, that “Student,” was a mathematically savvy chemist working as the chief brewer for the Guinness Corporation, and had to do a lot of the cranking himself. In the 1920s it took a lot of effort, in other words, to get an exact value of t for each degree of freedom. The high price did not stop the Significance Mistake from coming. But it came, slowly at first, and now like an avalanche. Eighty years on, it is essentially free to find a t . Credit the desk-top computer. “The cheapest way to get marketable results,” as Kruskal said.

To unblock the journal referees and editors, and break out of what Morris Altman calls “a steady-state low-level equilibrium,” we propose to try real market forces, as against the monopolized traditions of journal editing at present. If some auto mechanics used unsafe procedures to fix brakes—comforting themselves by saying that after all a young auto mechanic has to run his career—it would be a service to public safety and to the good mechanics to publicize the names of the good ones. Market forces would reward them. We propose an Angie’s List of economists who understand the insignificance of statistical significance. We will ask a carefully constructed set of major economists and econometricians (every editor of every major journal, for example) to state publicly and for publication by return-addressed postcard their support for the following propositions:

1. Economists should prefer confidence intervals to other methods of reporting sampling variance;

CORRESPONDENCE

2. Sampling variance is sometimes interesting, but a low value of it is not the same thing as scientific importance;
3. Economic significance is the chief scientific issue in economics; an arbitrary level of sampling significance is no substitute for it;
4. Fit is not a good all-purpose measure of scientific validity, and should be deemphasized in favor of inquiry into other measures of importance.

We will publish the names and responses, or lack of response, of everyone asked. On reflection any economist and econometrician who understands basic statistics will of course agree with the implied standards. In the decades since 1919 that we and scores of others have been making this point, no one has ever been able to *defend* the practice of significance testing. True, many people have gotten angry at the challenge. But no one has actually *met* it.

Stephen T. Ziliak
Roosevelt University

Deirdre N. McCloskey
University of Illinois-Chicago and
Erasmus University, the Netherlands

Editors,

Ziliak and McCloskey's indictment of economic studies misusing statistics ([EJW August 2004](#)) does not go far enough. A critical question is missing from their survey. Namely, does the study use out-of-sample data to test its model. This is not the same as testing against the correct null hypothesis, an issue ZM deal with.

Any new scientific theory (H1) has to go through two separate processes to be accepted as a replacement for the status quo (H0). First is its creation, second is its testing to see if it *predicts* better than H0.

A theory can be created from a mathematical derivation as in physics, purely from empirical data as in most other sciences, or some combination of both. In a statistical model, even if you have run all your tests and calculated your coefficients absolutely properly you have only accomplished the first step, creating H1. Getting a statistically significant result against the correct H0 merely suggests you are on the right track. To truly test and

verify H1 you need to test it against H0 with data *different* and *independent* from that from which H1 was created.

In econometrics this verification can only come from testing H1 on data not used to create it. While it would be sufficient to show that H1 was wrong because the author calculated wrongly in any of the many ways ZM list, this is not a necessary condition to reject H1.

An alternative to fitting coefficients is to stipulate them before doing any testing. Coefficient values could be stipulated by what makes economic sense. However, even in this case it's possible that some implicit data fitting is occurring because any practitioner is going to have some familiarity with the data ahead of testing just from becoming conversant in the field. Even if you are careful about implicit fitting in this case the ultimate test of H1 can still only occur once H1 is tested on sufficient data that occurs after H1 is created. This may seem like an overly cautious standard but it is the only way to guard against implicit fitting.

In practice if a model H1 is fitted on data from time period t_0 - t_1 (in-sample) it then needs to be tested on data from time period t_1 -present (out-of-sample). This doesn't mean refit H1 on out-of-sample data but test it to see if it predicts better than H0 in this time period. You can create an H1 with the highest t-values, R2, and what have you from in-sample data but it's completely meaningless until you see if H1 out predicts H0 in out-of-sample data.

In ZM's wonderful example of Milton Friedman's brief foray into metallurgy, the out-of-sample test of his theory was done by actually testing the newly created metal. In the physical sciences this concept boils down to can the experiment be reproduced. If and only if the experiment can be independently verified is H1 accepted. In 1989 when Martin Fleischmann and Stanley Pons reported their experiments which showed cold fusion, the scientific world did not immediately accept the implications. Certainly scientists were dubious because of the conflict with known theory, but it was also that the idea had to await additional independent experimental verification which never came.

Unlike physical sciences where you can create more data by running more experiments or in biology by getting more samples or subjects, in economics and finance you can only wait for the passage of time if you've

used all your data as in-sample data. Unfortunately there are no short cuts so the wise econometrician partitions the data into in-sample and out-of-sample before any hypothesizing.

Say in the case where you have been careful in segregating your data, creating an H1 that's both economically and statistically sensible and now when you run your test on out-of-sample data the model fails. At this point it is not good enough to go back to the drawing board and come up with some tweaks in your in-sample data and retest. Once the test is made on out-of-sample data the out-of-sample data now becomes part of the in-sample data for any future hypothesizing. Any additional testing must occur on new data. If you use the important ideas ZM discuss beforehand it makes it more likely H1 will prove out but going through these steps is no guarantee of success and certainly not sufficient to claim H1 has been verified.

In applied finance, good quantitative traders know these rules well. These traders also know not to play other games like selecting your in-sample data from a time period after your out-of-sample data. To ignore these rules means creating models that are likely to lose money once they are put into use and perhaps having to find a new line of work. Financial markets are very unforgiving.

For the work of economists to be as rigorous as other scientists it needs to be tested on data outside their samples used in their specific research. A new theory cannot be said to be explanatory until it is shown to predict better than alternatives. Fleischmann and Pons became pariahs in their fields not because they made an error in their experiments but because they acted as though their findings should be accepted before they were verified.

Bob Gelfond
CEO and Founder
MagiQ Technologies, Inc.

ZILIAK AND MCCLOSKEY RESPOND:

We agree with Robert Gelfond that a test on out-of-sample properties is a good argument. We do not agree, though, that it is necessary, a test that

“any new scientific theory must go through.” If this was what scientific progress meant there would be no progress in the historical sciences, in which the data are all in, such as evolutionary biology or geology or cosmology or history itself. Out-of-sample prediction, like a test on other sorts of data entirely or a consideration of logical coherence or a *gedankenexperiment*, is a useful option in scientific rhetoric. But there's no formula for scientific argument, and no timeless rules. We don't think that prediction is a rhetorical gold standard in science against which all arguments are to be tested. As we say: that way lies the end of geology and history.

We are sorry that Gelfond accepts the conventional calumnies on Fleischmann and Pons, and suggest that he have a look at a book by Eugene F. Mallove, *Fire From Ice: Searching for the Truth Behind Cold Fusion* (Wiley, 1991). We think it was thugs at Cal Tech and MIT, outraged that mere chemists would claim to have done some physics, that settled the issue in the minds of readers of *The New York Times*. But hundreds of articles are still published each year on the phenomenon Fleischmann and Pons discovered. *The Times* is not a good source for history of science.

Nor do we agree—and this is what really matters here—that “getting a statistically significant result against the correct H_0 suggests you are on the right track.” Our main point is that it suggests nothing of the kind. Gelfond speaks of “both economically and statistically sensible.” If by “sensible” he means “passing conventional levels of significance, modulo sample size,” he's missing our main point. It is: arbitrary levels of significance don't matter for science at all. The test of a financial model is its impact on a trader's bank balance—an economic criterion—not any statistic in itself. We note, alas, that this Good Old Chicago School point has recently given way in finance to more and more sophisticated—but financially and scientifically irrelevant—tests of “significance.”

So we worry then that Gelfond has not grasped our simple, central point. But that's not unusual. It's been our experience for years and years and years!

Econ Journal Watch welcomes letters commenting on the journal or articles therein. Please send correspondence to editor@econjournalwatch.org, Subject line: EJW Correspondence.